DOCUMENT RESUME

ED 351 345 TM 019 147

AUTHOR	Koretz, Daniel
TITLE	The Vermont Portfolio Assessment Program: Interim Report on Implementation and Impact, 1991-92 School Year. Project 3.2: Collaborative Development of Statewide Systems. Report of Year 1 Vermont Study.
INSTITUTION	National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.; Rand Corp., Santa Monica, Calif.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	92
CONTRACT	R117G10027
NOTE	83p.
PUB TYPE	Reports - Descriptive (141) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Data Collection; *Elementary School Teachers; Grade 4; Grade 8; Intermediate Grades; Junior High Schools; Mathematics Achievement; Pilot Projects; *Portfolios (Background Materials); Principals; Program Evaluation; Program Implementation; Questionnaires; *Secondary School Teachers; *State Programs; *Student Evaluation; Test Use; Writing Evaluation
IDENTIFIERS	Administrator Surveys; *Performance Based Evaluation; Teacher Surveys; *Vermont

ABSTRACT

Since 1988, Vermont has been developing a state assessment program, the centerpiece of which is the student portfolio and "best pieces" drawn from it. The state's use of portfolio results will be limited to reporting aggregates, but within schools, teachers and principals may use the portfolios as they choose. The assessment program, which encourages the adoption of higher standards as it coexists with Vermont's tradition of local control and innovation, currently includes mathematics and writing at grades 4 and 8, and will eventually encompass a broader range of subjects. Pilot implementations of these programs were carried out in the 1990-91 school year. This interim report describes preliminary results from 4 data collection efforts related to mathematics portfolios: (1) questionnaires sent to 40 fourth-grade and 40 eighth-grade mathematics teachers in pilot schools (18 percent response rate); (2) questionnaires completed by 8 mathematics portfolio raters in 1990-91; (3) teacher questionnaires from 172 teachers (83 percent response rate) in 1991-92; and (4) interviews with 77 principals in 1991-92. Teacher and administrators responses concerning the use of portfolios will be incorporated into later studies. Teachers reported several positive effects on instruction and learning. Three appendixes provide a description of the study samples, the year 2 principal interview and instructions, and the year 2 teacher questionnaire. (SLD)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED351345

National Center for Research on
Evaluation, Standards, and Student Testing

Final Deliverable – February 1992

Project 3.2: Collaborative Development
of Statewide Systems

Report of Year 1 Vermont Study

The Vermont Portfolio Assessment Program: Interim Report on Implementation and Impact, 1991-92 School Year

Daniel Koretz, Project Director
The RAND Corporation

U.S. Department of Education
Office of Educational Research and Improvement
Grant No. R117G10027 CFDA Catalog No. 84.117G

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

BEST COPY AVAILABLE

TM019147

The results reported here are preliminary and have not been reviewed. Final results of these and other studies of the Vermont assessment program will be published by RAND during the 1992-93 school year.

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education, and by a grant from the John T. and Catherine D. MacArthur Foundation.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

Table of Contents

Summary vii
Pilot-Year Rater Survey viii
1991-92 Principal Interviews ix
The Pilot-Year and 1991-92 Teacher Questionnaires ix
Acknowledgments xiii
Introduction 1
Background on the Vermont Assessment Program 2
Mathematics 4
Writing 7
CRESST/RAND Studies of the Vermont Program 8
Pilot Year Questionnaires 8
Teacher Survey 9
Rater Survey 14
Goals and Designs of the Second-Year Studies 16
Data Collection 17
Samples of Schools 17
Results of the Second-Year Principal Interviews 18
Use of Portfolios 20
Burdens Imposed and Support Needed 21
Perceived Attitudes of Teachers, Students, and Parents 25
Perceived Effects on Instruction 30
Perceived Effects on Traditional Mathematics Performance 32
Results of the Second-Year Teacher Questionnaire 33
Teacher Characteristics 33
Description of Participating Classes 34

Time Devoted to Portfolios 35
Training and Support 36
Changes in Curriculum and Instruction 38
Preparation and Revision of Portfolio Entries 42
Teachers' Evaluations of Student Performance 43
Effects of Portfolios 44
Portfolio-Related Problems 45
References 48
Appendix A: Description of the Study Samples 49
Appendix B: Year-2 Principal Interview and Instructions 55
Appendix C: Year-2 Teacher Questionnaire 63

List of Tables

Table 1: Teacher Time Spent on Portfolio Activities (Hours per month) 36

Table 2: Classroom Time Spent on Portfolio Activities (Hours per month) 37

Table 3: Teachers' Ratings of Training and Support Activities (% of Teachers) 39

Table 4: Change in Class Time Devoted to Mathematical Topics (% of Teachers) 40

Table 5: Frequency of Positive Effects (% of Teachers Reporting) 44

Table 6: Frequency of Portfolio-Related Problems (% of Teachers Reporting) 46

Table A1: Grade 4 Final Primary RAND Sample 53

Table A2: Grade 8 Final Primary RAND Sample 53

The Vermont Portfolio Assessment Program: Interim Report on Implementation and Impact, 1991-92 School Year

Daniel Koretz, Project Director
Brian Stecher and Edward Deibert

National Center for Research on
Evaluation, Standards, and Student Testing

The RAND Corporation

Summary

Since 1988, Vermont has been developing a cutting-edge state assessment program, the centerpiece of which are student portfolios and "best pieces" drawn from them. The state's use of the portfolio results will be limited to reporting aggregates—statewide results and perhaps the results at the level of districts and schools. Within schools, however, teachers and principals may make use of the results for individual students as they choose. The assessment program was designed to serve many goals: to provide rich data on student performance; to encourage better teaching and the adaptation of higher standards; to co-exist with Vermont's strong tradition of local control and innovation; and to encourage greater equity of educational opportunity. The assessment program currently includes mathematics and writing in grades 4 and 8 and will eventually encompass a broader range of subject areas.

Pilot implementations of the mathematics and writing assessments were conducted in a sample of schools in the 1990-91 school year, and the first statewide implementation was carried out in 1991-92. These terms, however, may be misleading. From the outset, the Vermont Department of Education anticipated a long and decentralized development process in which committees of teachers, with the help of outside consultants and substantial trial and error, would gradually build the assessment program. Thus, particularly in mathematics, the "pilot" effort was not a true pilot in the traditional sense of a field test of an already designed program. Rather, the pilot was a developmental effort. Indeed, some of the key aspects of the program, such as important details about the scoring of pieces, were still being worked out during the statewide implementation in 1991-92.

Beginning in August 1988, RAND has consulted with Vermont about the development and eventual evaluation of the assessment program. Since 1990, RAND, as part of the Center for Research on Evaluation, Standards, and Student Testing (CRESST), has been carrying out a multi-faceted evaluation of

the assessment program and its effects. Because of the diversity of the program's goals and the pervasiveness of the changes it requires, the CRESST/RAND evaluation is broad in scope. It will address a variety of questions pertaining to the implementation and operation of the program at the school and classroom level, the quality of measurement provided by the system, and its effects on instruction and other aspects of schooling.

This interim report describes the results of four data collection efforts: questionnaires administered to participating teachers and mathematics portfolio raters during the 1990-91 pilot year, and teacher questionnaires and interviews with principals administered during the first statewide implementation in 1991-92. *The information in this report is preliminary and has not been reviewed;* it will be integrated with additional information in a later report for public distribution.

Pilot-Year Rater Survey

In the 1990 pilot, all mathematics portfolios were scored by a single group of eight specially trained raters. All eight responded to a detailed questionnaire about their experiences rating the portfolios. Given the small number of raters and the developmental nature of the pilot implementation, the raters' responses do not constitute an evaluation of the Vermont program as implemented statewide in 1991-92. However, they highlight a number of issues that arose in developing the program and that will confront portfolio-based assessments more generally, and perhaps some other performance-based assessments as well.

Raters had mixed reactions when asked about the scoring process. They were required to score each piece on a 4-point scale for each of seven criteria and then to create an aggregate rating on each criterion. Grade-8 raters reported relatively little difficulty scoring the individual pieces, but all grade-4 raters found it difficult in a substantial number of cases. One reason for this difficulty was that raters found that some pieces did not match all the scoring criteria well. (Some teachers who were unfamiliar with the program sent inappropriate materials, such as computation tasks, but to some extent this problem was a reflection of the wide latitude in selection of appropriate tasks permitted by the portfolio approach.) Another problem was that many pieces were not accompanied by sufficient documentation of the original assignment, making it difficult to understand what the students' performance had been. Providing aggregate scores for each student's set of tasks proved more difficult; the scale points that had been devised to be appropriate for individual tasks were inappropriate for averaging. (The procedure for aggregating was accordingly changed in 1991-92.) A number of raters also noted that the volume of scoring they were required to undertake was burdensome.

Raters were concerned about the nature of the tasks students were given as well. Most of the raters believed that variations in the tasks assigned to students significantly influenced ratings in a substantial number of cases.

For example, they believed that some students had been given tasks that afforded insufficient opportunity to display the desired aspects of performance.

1991-92 Principal Interviews

During the second half of the 1991-92 statewide implementation, we interviewed a representative sample of 77 principals about their experiences with the portfolio program. All but three were new to the program; three had participated in the 1990 pilot.

Although the responses of the principals were diverse, a clear pattern emerged: the typical principal perceived the program as a worthwhile burden.

A large majority of the principals said that the program was burdensome, but they often pointed to burdens on participating teachers, not on them as administrators. The burdens they cited were diverse: difficulties with the state's administration of the program, difficulties adapting to what many saw as a rapid pace of implementation, the burdens of record-keeping and logistics, and the overall time demands of the program. The great majority of principals had provided release time to teachers to attend training sessions. A substantial number of principals considered the burden to be partly a transient problem associated with the novelty of the program.

Despite this perceived burden, many of the principals' comments on the program were clearly positive. A minority of principals had no opinion yet about the impact of the program on mathematics instruction, but almost all the majority who offered a view were satisfied that its effects have been positive. Almost none of the principals characterized the attitudes of their teachers toward the program as predominately negative; about a fourth characterized their attitudes as positive, and nearly three-fourths characterized the attitudes of their teachers as mixed. Principals offered a long and diverse list of attributes of the program that sparked positive responses by their teachers; many cited its effects on curriculum and instruction, while others noted its value for within-class and within-school assessment purposes. Perhaps the most telling sign of support for the portfolio program is the finding that it had already been extended beyond the grades targeted by the state in about half of the schools. Another 13% of the principals said they hope to extend the program beyond the target grades next year, and a few indicated that they are encouraging other teachers to keep portfolios. Indeed, 9% of the principals reported that *all* of the teachers in their schools are already keeping portfolios.

The Pilot-Year and 1991-92 Teacher Questionnaires

In 1990-91, we mailed questionnaires to all teachers participating in the mathematics pilot; in 1991-92, we sent a partially overlapping questionnaire to a sizable sample of teachers statewide. The questionnaires were designed to gather information about the implementation of the portfolio assessment and its effects on teachers and students.

The response rate to the pilot-year teacher questionnaire was low, so the results are only suggestive. They suggest that the portfolio assessment was still in a very formative stage. Although the overall goals of the program appeared to be clear to teachers, procedures for implementing the portfolios were not. For example, many teachers felt that training was insufficient because it failed to provide specific guidelines for implementing the portfolios, adequate numbers of examples of appropriate activities, and clarification of the criteria to be used to judge the portfolios.

Nevertheless, portfolios had an impact on instruction. The greatest effect appeared to be increasing the amount of attention paid to mathematical communication, particularly writing about mathematics.

Participating teachers raised concerns about the interpretation of portfolio scores. On the one hand, teachers felt the portfolios demonstrated students' logical thinking and problem-solving abilities better than standardized tests. On the other hand, there was concern that the work in some students' portfolios did not demonstrate the students' true ability because of lack of interest in the tasks, because the students had the option to revise the product repeatedly, and because of the confounding of mathematical ability with writing ability (particularly at grade 4).

During the second year (1991-92) of the study, the questionnaire response rate was high (83%), and the results are more representative of Vermont fourth- and eighth-grade teachers as a whole. They indicate that the portfolio assessment was implemented more smoothly, that training was more widely available and more effective, that the assessment continued to have significant impact on students and teachers, but that some difficulties remained.

Virtually all the fourth- and eighth-grade teachers received some type of formal, state-sponsored training in the implementation of the mathematics portfolios, and approximately three-quarters reported that the training prepared them adequately well to work with the math portfolios. Still, most teachers also reported occasional difficulties knowing what they were expected to do or how to do it. This suggests a continuing need for training and support in some form.

Teachers devoted considerable effort to the portfolios, both in class and outside class. On average, teachers spent six hours per week working on portfolios: two to three hours per week preparing for portfolios, a similar amount of time in classroom portfolio activities, and about one hour per week scoring or evaluating portfolios. Students, too, devoted a substantial amount of time to portfolio activities. Teachers reported that on average, students spent about three and one-half hours per week on portfolios: about two hours doing portfolio tasks, about one hour per week revising them and about one-half hour per week organizing their portfolios. Fourth-grade students spent significantly more time doing and revising tasks than eighth-grade students.

Consistent with the goals of the program, teachers at both grade levels reported marked increases in the time devoted to teaching the topics of problem solving and patterns/relationships. In contrast, there were only limited changes in the amount of time devoted to computation, estimation, and measurement/geometry. Similarly, there was an overall increase in the emphasis teachers placed on mathematical communication activities such as describing personal experiences relating to mathematics; having oral discussion of mathematical topics; making charts, graphs and diagrams; and having students write reports about mathematics. These changes were more pronounced for fourth-grade teachers than for eighth-grade teachers.

Because of the flexible nature of the mathematics portfolio assessment and the lack of strict guidelines for the composition of portfolios and the preparation of individual pieces, teachers exercised considerable discretion over the work produced by their students. One result of this freedom was that mathematics portfolios were not implemented in the same way by all teachers. Teachers established different rules for revision and collaboration, and they emphasized different aspects of quality in the final work products. Teachers reported that student work was revised once, on average, before being placed in students' portfolios.

Teachers reported that the mathematics portfolios had a number of positive effects on themselves and their students. For example, more than one-half of the teachers said they were frequently more enthusiastic about teaching math, and over 90% were more enthusiastic at least occasionally. Perhaps more significantly, more than 80% of the teachers indicated that they had changed their opinion of students' mathematical abilities based upon the students' portfolio work. Although the amount of reported change was small, the fact that such a large proportion of teachers saw student performance that changed their estimates of students' abilities seems significant.

Time continued to be the major source of problems. In the view of teachers, the greatest problems created by the portfolios were not about what to do, but when to do it. Finding time to cover the full mathematics curriculum was the most widespread problem, followed closely by finding time to prepare portfolio lessons.

Teaching students to solve problems was the other major concern. Although teachers perceived students to be interested in the portfolio tasks, they reported that students frequently had difficulty solving portfolio problems. One-half of the fourth-grade teachers and over one-third of the eighth-grade teachers said students often did not know how to solve problems.

Finally, teachers perceived portfolios as having both positive and negative effects on the performance of low ability students. Approximately one-quarter of the teachers reported that low ability students were frequently more successful as a result of the portfolios. Others, however, reported that the portfolios created greater problems for low ability students.

Acknowledgments

The work reported here required the assistance of many people. We would like to thank the more than 200 Vermont teachers and principals who, despite the time demands of implementing a major reform, graciously contributed their time to make this study possible. Eric Hamilton tabulated the teacher questionnaire results and helped draft the description of them. Mary Ann Minardo conducted a substantial share of the principal interviews reported here. We are grateful to Advanced Systems for Measurement in Education—and in particular, Mark Wiley—for distributing the Year-2 teacher questionnaires and for providing us the responses in electronic form.

THE VERMONT PORTFOLIO ASSESSMENT PROGRAM: INTERIM REPORT ON IMPLEMENTATION AND IMPACT, 1991-92 SCHOOL YEAR

Daniel Koretz, Project Director
Brian Stecher and Edward Deibert

National Center for Research on Evaluation, Standards, and Student Testing

The RAND Corporation

Introduction

Since 1988, Vermont has been developing an assessment program that is at the forefront of innovation in large-scale assessments. Although rapidly growing numbers of statewide assessment programs incorporate some form of performance assessment, the Vermont program is unusual among them in that its centerpiece is student portfolios and "best pieces" drawn from them. Portfolio-based assessment is not new, but Vermont was first state to make portfolios the backbone of a statewide assessment program. The Vermont program is also unusual in the degree to which it is "bottom up": many aspects of the assessments in each subject are worked out iteratively by committees of teachers, and classroom teachers retain wide latitude in implementing the program.

The Vermont Department of Education selected 48 schools to pilot fourth- and eighth-grade assessments in writing and mathematics during the 1990-91 school year, and 90 other schools asked to participate to varying degrees in the pilot efforts that year. The first statewide implementation of the assessment in those two grades and subjects was conducted in the 1991-92 school year.

Beginning in August 1988, RAND has consulted with Vermont about the development and eventual evaluation of the assessment program. Since 1990, RAND, as part of the Center for Research on Evaluation, Standards, and

Student Testing (CRESST), has been carrying out a multi-faceted evaluation of the assessment program and its effects.

This interim report provides background on the Vermont Assessment Program, describes the goals and components of the CRESST/RAND study, and discusses the results of four data collection efforts: questionnaires administered to participating teachers and mathematics portfolio raters during the 1990-91 pilot year, and teacher questionnaires and interviews with principals administered during the first statewide implementation in 1991-92. *The information in this report is preliminary and has not been reviewed.* The information provided here will be reviewed and integrated with information from a variety of other sources in a full report of the 1991-92 CRESST/RAND study that will be issued during the 1992-93 school year.

Background on the Vermont Assessment Program

Until recently, Vermont had no regular statewide assessment program. By the late 1980s, however, pressure was building to provide regular information on student performance, and by 1988, the state Department of Education began movement toward establishment of a statewide assessment system.

The deliberations that led to the decision to build the present, portfolio-based system are difficult to summarize succinctly because they were lengthy and involved many diverse people, including the Commissioner of Education (Rick Mills), the Department's then-Director of Policy and Planning (Ross Brewer), the governor, members of the state board, local board members, teachers, and others. Several persistent themes, however, were stressed by Mills, Brewer, and others working to build the system.[1] Ideally, the new system would:

- Avoid the distortions of educational practice that conventional test-based accountability appeared to have created in some other states;

[1] This description is based in large part on the first author's participation in meetings and discussions with Department of Education staff and others involved in building the assessment program. No single source summarizes the development of the program, but many of the points noted here have been described elsewhere. See, for example, Vermont Department of Education, 1990, 1991a, 1991b, and Mills and Brewer, 1988.

- Encourage good practice and be integrally related to the professional development of educators;
- Reflect the Vermont tradition of local autonomy, "encourage local inventiveness, [and] preserve local variations in curriculum and approach to teaching" (Mills & Brewer, 1988, pp. 3, 5);
- Provide "a high common standard of achievement for all students" (Mills & Brewer, 1988, p. 3); and
- Encourage greater equity in educational opportunity.

Those responsible for the nascent program were aware of the difficulties inherent in having an assessment program serve many functions at once and had been warned that some of their goals for the program pointed to different assessment designs. For example, a system designed to provide rich information about students and positive incentives for teachers might look very different from a system that was designed primarily to provide highly comparable information across schools.[2] The system that eventually emerged was intended to be a compromise among the many goals for the system; for example, it should provide reasonable comparability across schools, but not at the cost of stifling good practice and local innovation.

The basic outline of the assessment program emerged quite quickly. Eventually, the assessment would span a broad range of subjects, but the state decided to begin with assessments in writing and mathematics in grades 4 and 8. The assessment would have three components: year-long student portfolios, "best pieces" drawn from the portfolios, and state-sponsored "uniform tests." The state would use the results for reporting aggregate patterns of performance (statewide and, it was hoped, at the level of districts and schools). School staff, however, could use the results, including the scores for individual students, in whatever ways they chose.

The details of the program, however, have been worked out only gradually. In contrast to the many states that either buy off-the-shelf tests or contract to have new tests built on a short schedule, the Vermont program was seen from the outset as a long-term and decentralized development effort. For example, in 1988, Mills called for mixing state-of-the-art assessment techniques with "emerging" techniques and warned that the development of

[2] Dan Koretz, presentation to Commissioner Mills, Governor Kunin, and others, August, 1988.

the new program would be "a very long effort" (Mills, 1988). Thus, in both subjects, the so-called "pilot" implementation in 1990-91 was less a true pilot of a developed program than an integral part of the development effort. Indeed, in mathematics, even the first full statewide implementation in the 1991-92 school year would be most accurately categorized as a combination of a developmental effort and a pilot test, rather than as an initial implementation of a fully planned program. Some of the scoring details of best pieces in the 1991-92 statewide implementation, for example, were not resolved until spring of 1992, and procedures for rating entire portfolios remain under development.

Primary responsibility for the development of the portfolio and best-pieces components of the program was given to state-sponsored committees of teachers. These committees worked independently of each other, so the program evolved differently in writing and mathematics.

Mathematics

As implemented in 1991-92, the mathematics program required that students and teachers cull from each student's portfolio a set of five to seven "best pieces." Teachers were requested to include among the best pieces exemplars of three types of problems: puzzles, investigations, and applications. According to the *Teacher's Guide*,[3] *puzzles* are tasks that "require students to identify and explore approaches to non-routine problems ... [where] most of the problem deals with logic and reasoning." For example, "With a seven minute hourglass and an eleven minute hourglass, how could you time the boiling of an egg for 15 minutes?" *Applications*, by comparison, "require students to apply knowledge they already possess." For example, "A mature tree can utilize 13 lbs. of carbon dioxide in a year. The average car spews out 4000 lbs. of carbon dioxide in a year. How many mature trees would you need in order to use this much carbon dioxide?" (The problem continues, exploring the relationship of fuel efficiency and carbon dioxide emissions.) *Investigations* include "explorations, data collection and analysis or some level of research that leads to conclusions." For example, "Given a piece of graph paper, determine the size of a square to be cut out of each corner of the graph paper that will allow you to fold the graph paper into an open box with the

[3] *Vermont Mathematics Portfolio Project: Teacher's Guide*, Vermont Department of Education, September, 1991d.

greatest volume. Try squares from 1 to 9 units. Then determine the volume of each open box after you have folded the paper. Collect the data in a data table of your design. Graph the results. Write up your results." The *Resource Book* provided to each teacher contained many samples of each type of task.[4]

The best-pieces sets of a sample of students from each participating classroom were sent to one of seven central sites for scoring by groups of volunteer teachers. (The samples from each class were selected by the state's assessment contractor, not the classroom teachers.) Training in scoring had been offered statewide before the scoring sessions, but the volunteer raters varied substantially in their level of training, and supplementary training was provided at the beginning of each three-day scoring session. (Beginning in the 1992-93 school year, the state plans to switch to a system in which teachers score their own portfolios and then submit a sample for moderation. That system is already in use in the writing assessment and is described below.)

All the best pieces were graded on 4-point scales against seven criteria, four classified as problem-solving and three as communication. The seven criteria, and the scale points on each for individual pieces, were as follows:

- Understanding the task: 1. totally misunderstood; 2. partially understood; 3. understood; 4. generalized, applied, extended.
- How—approaches/procedures: 1. inappropriate or unworkable; 2. appropriate some of the time; 3. workable; 4. efficient or sophisticated.
- Why—decisions along the way: 1. no evidence; 2. possible; 3. inferred with certainty; 4. shown/explicated.
- What—outcomes of activities: 1. no extension; 2. observations; 3. connections, applications; 4. synthesis, generalization, abstraction.
- Language: 1. no or inappropriate; 2. some of the time; 3. most of the time; 4. rich, precise, elegant.
- Representation: 1. no use; 2. use; 3. accurate and appropriate; 4. perceptive.
- Presentation: 1. not clear; 2. some clear parts; 3. mostly clear; 4. clear.

[4] *Vermont Mathematics Portfolio Project: Resource Book*, Vermont Department of Education, September, 1991c. The guide, which was prepared by a committee of teachers and distributed by the state Education Department, contains a collection of activities appropriate for inclusion in mathematics portfolios.

The ratings on the individual pieces were then aggregated to provide an overall rating of the entire set of best pieces on each of the seven criteria. However, the nature of the scales for individual pieces precluded a simple averaging across pieces. A set of three 4s ("perceptive") and three 2s ("use") on "Representation" do not average to a 3 ("accurate and appropriate"). Accordingly, the mathematics committee developed a "holistic" rating system in which the 4 scale points for the aggregate rating were defined differently than those for the individual ratings, and a fixed algorithm (based on the number of instances of different scores on the individual pieces) was used to create the aggregate scores.

This system for aggregating scores was too detailed to warrant a full description here, but a brief description of the ratings on the "Understanding the task" criterion illustrate the general approach. The four scale points for the holistic rating of the set of best pieces were:

1. Totally misunderstood in more than half of the best pieces.
2. May have understood or read beyond the surface problem in some instances but only partially understood more than half of the time or totally misunderstood the problem in two or more instances.
3. Understood or read beyond the surface problem most of the time but partially understood or misunderstood in some instances.
4. Understood the task most of the time and read beyond the surface problem at least a couple of times.

Each of these descriptions was accompanied by as many as 10 combinations of scores from individual pieces that would produce the aggregate rating. Only the five best pieces were counted, regardless of whether the student included one or two more (as guidelines permitted).

The mathematics portfolio assessment was accompanied in the 1991-92 statewide implementation by a sample-based administration of the state's new Uniform Test (UT) of mathematics. The UT is a matrix-sampled, mixed-format test, combining multiple-choice and open-ended items. Unlike the portfolio assessment, the UT was designed and scored by Advanced Systems for Measurement in Education, Vermont's testing contractor.

Writing

The design of the writing assessment, which was largely completed during the 1990-91 pilot year, is substantially different from that of the mathematics assessment. In writing, students' portfolios must include a set number of pieces of specified types, the entire portfolio is rated, and a single best piece is chosen.[5]

In grade 4, each students' portfolio must include:

1. A table of contents.
2. A single best piece, which is selected by the student, can come from any class, and need not address an academic subject.
3. A letter explaining the composition and selection of the best piece.
4. A poem, short story, or personal narration.
5. A personal response to a book, event, current issue, mathematical problem, or scientific phenomenon.
6. A prose piece from any subject area other than English or language arts.

The requirements for eighth grade are the same except that the portfolio must include three prose pieces.

The best piece, the rest of the portfolio, and performance on the Uniform Test of writing (which is a direct writing task using standardized conditions and a single prompt) were all scored on the same five dimensions:

- Purpose;
- Organization;
- Details;
- Voice/Tone; and
- Usage/Mechanics/Grammar.

A single 4-point scale is used with all five criteria, labeled as: 1. rarely; 2. sometimes; 3. frequently; 4. extensively. The descriptions of the scale points, however, are generally phrased in terms of *quality* or *extensiveness,* not

[5] The following description is taken largely from *"This is my Best:" Vermont's Writing Assessment Program, Pilot Year 1990-91*, Vermont Department of Education, undated.

frequency. For example, in the case of purpose, the description of "sometimes" is:

- "Attempts to establish a purpose.
- Demonstrates some awareness of audience and task.
- Exhibits rudimentary development of ideas."

(Vermont Department of Education, 1991b, p. 6)

In the 1991-92 statewide implementation, teachers scored their own students' portfolios and best pieces. Advanced Systems scored the Uniform Tests and arranged for a sample of portfolios to be drawn from each class for "moderation"—that is, to be scored by an external panel of teachers so that the scores of participating teachers could be calibrated to a common standard.

CRESST/RAND Studies of the Vermont Program

The characteristics of the Vermont assessment program require that the CRESST/RAND evaluation be broad in scope. The CRESST/RAND evaluation is a series of interrelated efforts designed to gather information about:

- The implementation and operation of the program at the school and classroom level;
- The quality of measurement (including reliability and validity); and
- Effects on instruction and on other aspects of schooling.

Our efforts initially have been focused primarily on mathematics because of the novelty of the mathematics program. Those designing the writing assessment have been able to draw on substantial past experience with direct assessments in other locations. In contrast, those building the mathematics program were plowing new ground; they did not begin with a consensus about the attributes that should be measured and found a paucity of past experience upon which to draw.

Pilot Year Questionnaires

Although most of the CRESST/RAND work during the 1990-91 school year was developmental, two preliminary surveys were fielded. One was a survey of mathematics teachers in the sampled pilot schools. The second was a

questionnaire administered to the small group of teachers (four in each grade) who scored all of the pilot-year mathematics portfolios.

The results of the two questionnaires must be interpreted in the light of the developmental nature of the 1990-91 "pilot." As noted above, this was not a pilot in the true sense of the word—that is, a test of a largely planned and designed program. Many of the key elements of the program had not been designed or planned when the pilot began in 1990, particularly in mathematics. Thus, it would be more accurate to characterize the 1990-91 work as a developmental and experimental effort, a substantial part of which was devoted to trial and error and redesign.

The responses of teachers and raters to our questionnaires reflect this and should not be seen as a summary evaluation of the program. Nonetheless, the responses provided a number of suggestive findings that were useful for formative purposes (both in shaping the assessment program itself and in finalizing the CRESST/RAND design for 1991). For example, the criteria used to judge mathematics portfolios have since been clarified, the method used for aggregating mathematics ratings has been changed, and a substantial number of illustrative tasks and benchmarks for scoring have been provided—all changes that address some of the difficulties that appeared in the questionnaire responses. In addition, the responses illustrate a variety of issues that will arise in a wide variety of performance-assessment programs.

Teacher Survey

The survey of teachers was in the form of a mail questionnaire. All mathematics teachers in the pilot schools were sent questionnaires, but the response rate was poor, about 18%. (On the positive side, the teachers who responded to the survey did not appear to be dramatically different than the population of teachers in Vermont in terms of experience, classroom organization, school size, etc.) Data collection strategies were revised for 1991-92 in the light of this unacceptable response rate. Although the poor response rate, coupled with the poorly defined nature of the program itself during the pilot year, make it impossible to draw firm conclusions, the questionnaire did yield suggestive and useful findings.

Goals of the program. Teachers seemed to understand, in general terms, the purposes of the portfolios and the types of student work that were to be

included. Teachers' descriptions of the work to be included in student mathematics portfolios focused on demonstrations of problem solving or higher order thinking skills. Many also mentioned work involving writing about mathematics and using mathematics in other disciplines, as well as pieces that demonstrated growth or progress in students' mathematical understanding. There was widespread understanding that computational exercises, worksheets, tests, etc. should not be included in the portfolios.

However, a few teachers (primarily at the fourth-grade level) seemed to think that almost any non-traditional mathematics-related work was appropriate, for example, "photographs of group work," "puzzles," "artwork that shows mathematics in use," "creative mathematics" (hands-on projects), and "all things fun and creative that have been enjoyed by students."

Staff development and support. Once the state-sponsored, regional workshops were completed, teachers received little, if any, additional support in implementing the mathematics portfolios. All of the fourth-grade teachers and approximately three-quarters of the eighth-grade teachers attended the statewide orientation meeting held in September 1990, and at least one regional workshop held later in the school year. In fact, most attended two or three regional workshops in mathematics (and most of the fourth-grade teachers attended one or two additional workshops in writing). Thus, teachers who responded to the survey received most of the information available from the state regarding portfolios. Furthermore, they reported that most of the other fourth-/eighth-grade teachers in their schools attended one or more of the state-sponsored workshops, as well. Approximately one-half of the principals attended one of the workshops.

Many teachers were dissatisfied with the state-sponsored workshops because they failed to provide specific guidelines for implementing the portfolios, adequate numbers of examples of appropriate activities, and clarification of the criteria to be used to judge the portfolios. Teachers who felt they received these specific kinds of information were pleased with the workshops; those who did not were dissatisfied. (The 1991-92 teacher questionnaire, reported below, revealed substantially higher levels of satisfaction with the training provided in the 1991-92 school year.)

Although there was some initial orientation and sharing of ideas among staff at local schools, it was limited in scope and duration. More than half of the fourth-grade teachers participated in a least one local school meeting to discuss the portfolios, but the total amount of time devoted to these discussions was typically only two or three hours. After these initial discussions, many teachers received little support from their peers and administrators. They "felt very isolated during the whole year." More than one-half of the teachers in both grade levels received no support from other teachers at their schools. Teachers in some schools shared materials and activities, and some even met regularly to help each other with "creative problem solving" ideas. However, a large number were operating on their own once the state workshops were over.

Portfolio practices. Half of the eighth-grade teachers used the portfolios with all five of their classes; half used them only with one or two classes. Most teachers began collecting students' work in the fall, soon after attending the state workshops. (One fourth-grade teacher waited until April, and one never used the portfolios at all.) The typical student's portfolio included approximately six pieces of work (regardless of grade level), but the number of pieces ranged from zero to 20.

Best pieces. There was considerable variety in how teachers implemented the best pieces component of the portfolios. To some degree, this reflected the unformed nature of the program in the "pilot" year. For example, about one-quarter of the classes did not select best pieces, often because of a lack of clarity about what was expected. In most classes the students and teachers selected the best pieces together. In some cases students made the selection on their own, and in two cases the teachers selected the best pieces. The selection of best pieces was guided by different criteria, including assignments on which students "learned the most," assignments students "enjoyed working on the most," "culminating unit activities," and "pieces they put the most effort into and which would impress the examiner the most."

Problems implementing portfolios. The most significant problem was the time required to produce work for the portfolios, time that had to be taken away from other aspects of the curriculum. The most common problems teachers reported concerned the mechanics of producing and managing portfolios. Chief among these problems was lack of time for working on portfolio entries. Both fourth- and eighth-grade teachers complained about the lack of time. For

example, one fourth-grade teacher responded, "I have no idea how much time to spend on the selection and 'rewriting' of best pieces. My students take an inordinate long time to rewrite. Time that can be spent on other math activities, I begrudge the time to be spent on rewriting [sic]." An eighth-grade teacher was concerned about covering the material in the curriculum: "My algebra and pre-algebra classes cannot afford to spend 20 class periods on portfolio work. The problem with assigning portfolio work for homework is that then it becomes collaborative efforts instead of individual work."

Teachers also voiced concerns about managing the paperwork associated with the portfolios. "Sorting and filing papers for all classes was a nightmare! Secretarial help would have been nice." Other procedural problems included:

- "Students being absent on a day we did a project."
- "Space in my classroom to collect all of the projects."
- "In doing group work, deciding whose portfolio gets the end product and how does the evaluator know who else worked on the project."
- "Inconsistent attendance and homework ... meant that many portfolios were incomplete."
- "Students wanted to include high test grades even if they were one-answer type tests."
- "Number of pieces: If an activity is to show growth, or lack of it, how many pieces should be included?"

(Data on the allocation of teachers' time to various aspects of the portfolio program during the 1991-92 statewide implementation are provided in the discussion below of the 1991 teacher questionnaire.)

Effects on curriculum. The greatest effect of the portfolio pilot on curriculum appeared to be increasing the amount of attention paid to mathematical communication, particularly writing about mathematics. Almost all fourth-grade teachers said they placed greater emphasis on writing in mathematics as a result of the portfolio pilot. More attention to writing also was the most commonly reported curricular change in eighth grade, but fewer than one-half of the eighth-grade teachers reported making any changes at all. A few of the teachers also indicated that they placed greater emphasis on explanations of students' mathematical thinking. These results are largely consistent with findings from the 1991-92 statewide implementation.

Students' reactions. Teachers described three types of student reactions to the mathematics portfolio pilot with approximately equal frequency: general dislike, initial fear or nervousness giving way to qualified enthusiasm, and enjoyment. In a number of cases, eighth-grade students' attitudes were linked to their concerns about grading (and they warmed when they learned they would not be graded on the portfolios). Although the numbers were far too small and the reactions too mixed to draw any strong conclusions, there appeared to be some relationship between teachers' perceptions of students' attitudes and teachers' reports of problems implementing the portfolios. (As noted below, a number of the principals we interviewed in the 1991-92 school year spontaneously commented that students' and teachers' attitudes toward the program appear to be linked.)

Drawing inferences about students' accomplishments. Teachers felt that the portfolios portrayed some student accomplishments better than standardized tests but also created some misconceptions about students' abilities. The strength of the portfolio model was its capacity to display students' creativity and individuality. This was evident both in students' choice of works ("one's trash is another's treasure") and in the manner in which they solved problems ("Some students are verbally and writing expression (sic) excellent thinkers and problem solvers"). The most common theme in the teachers' comments was that a person was more likely to obtain differentiated impressions of individual personalities in the portfolios than in standardized assessment.

Teachers also felt the portfolios demonstrated students' logical thinking and problem solving abilities (or lack thereof) better than standardized tests. By reviewing student portfolios, one might see "concepts being used appropriately in higher thinking situations;" "[students who] follow steps to problem solving and sometimes just get the answer wrong in computation mistakes;" "many portfolio entries indicate students' thinking processes and their learning process or the lack of such;" "students can apply math better than 'applications' subtests indicate."

Conversely, teachers noted a number of misconceptions about student accomplishments one might form by looking solely at student mathematics portfolios. The major concern was that the work in some students' portfolios did not demonstrate the students' true abilities. Comments included:

- "Some of my brightest students did not show interest, or do a good job, while some of my less able students did a tremendous job. Looking at the portfolios will give a different picture than my grade book."
- "[It appears] that the students have mastery of some skills that I don't feel they do. What's deceiving is that corrections were made and sometimes remade. This process is not included in the portfolio."
- "A few students are quite good in mathematics and think logically but cannot express themselves in writing as well. Simply reading their pieces may not give a true picture of their math skills."
- "Some students are disorganized and ... they had lost papers. Therefore, samples are sometimes not indicative of students' work/growth."

Fourth-grade teachers were more likely to be concerned about writing ability confounding mathematical ability in the mathematics portfolios. "Those students who are more mature and organized in their writing may appear to be the best math students. This is not always so. Age development, fine motor skills and organization can make a student appear in a portfolio to be more or less advanced than the student really is." "A few students are quite good in mathematics and think logically but cannot express themselves in writing as well. Simply reading their pieces may not give a true picture of their math skills." One teacher also noted that students often worked in groups and an individual student's portfolio entry may not have been his or her own work.

Rater Survey

In the 1990-91 pilot year, mathematics portfolios were formally scored only by a single group of eight raters, all of whom were specially trained teachers. All eight completed a questionnaire concerning their experiences in the rating process.

Classifying tasks. The system requires students to submit three types of tasks: puzzles, investigations, and applications. Raters reported little difficulty classifying tasks when asked how frequently they had difficulties, but they nonetheless pointed out several problems when asked for specific instances of tasks that were difficult to classify. The specific problems were of two types:

- Tasks that were by their nature difficult to fit into the three categories; and
- Tasks that were hard to classify because of insufficient information.

Ease of assigning a numerical score. Grade-8 raters generally found it easy to assign a 1-through-4 numerical rating to individual entries, but all grade-4 raters found it hard in "a substantial number" or "more than half" of the entries. The difficulties encountered by grade-4 raters were not confined to particular rating criteria or types of tasks.

Insufficient credit. We asked raters whether there were some tasks that deserved credit in their opinion for some aspect of performance that was not captured by the scoring criteria. Raters generally did not find this to be a serious problem, but a few comments suggest that puzzles and *insufficient documentation of student's work* may pose a problem in this regard.

Adequacy of final portfolio ratings. Raters were asked to summarize each portfolio for each of seven scoring criteria across all tasks included in the portfolio. Raters were less confident in these summary ratings than in their ratings of individual entries. Subsequent discussion revealed that part of the problem was that the scales were not defined to be interval: two tasks rated as 4 and two at 2 did not average to a 3. This aspect of the scoring procedure was accordingly altered for the 1991-92 statewide implementation, as noted in the description of the mathematics assessment above.

Credit for effort. Raters were not explicitly asked to rate students' level of effort, but we asked how adequately final ratings reflected differences in effort. Grade-8 raters were sanguine in their answers to this question, but their answers to the earlier questions point to substantial problems in discerning what kids actually *did* to solve a problem, which implies a lack of information on difficulty. Grade 4 raters were more negative. Two of the four said that this was a problem in a substantial number (but fewer than one-quarter) of the cases. One said that effort was difficult to gauge.

Adequacy of description of student assignments. Raters were asked whether the descriptions of assignments were sufficient to understand what had been required of students and to apply ratings accordingly. Raters were negative about this: Five said that information was inadequate in a substantial number (but fewer than one-half) of cases, and three said it was inadequate in

more than one-half. Moreover, most raters maintained that the nature of the assignments and thus the opportunities that students had to show skill levels significantly influenced ratings in a substantial number of cases.

Least successful elements of portfolio scoring. When asked specifically to list the two or three aspects of scoring that were least successful, raters pointed to a variety of things, some of which reflected the issues discussed above, including:

- Difficulty in scoring so many pieces.
- Difficulties with specific criteria: why, math language, and clarity. (One person questioned the necessity of the last.)
- Difficulties with the nature or presentation of tasks:
  - Tasks that provided insufficient opportunities for the desired performance.
  - Tasks that were unscorable or that focused on undesired performance (such as computation).
  - Entries that were not accompanied by the original assignment.

Only a single rater responded that inappropriate levels of difficulty of assigned tasks was a problem with scoring. However, a number of people pointed to that problem in subsequent conversations.

Goals and Designs of the Second-Year Studies

The CRESST/RAND efforts during the 1991-92 school year obtained data on diverse aspects of the basic questions of implementation, measurement quality, and impact. Among the implementation topics addressed in Year 2 were the burdens perceived by school staff, the time allocated to portfolio-related activities, the types of support requested by and provided to teachers, the perceived adequacy of state-sponsored training, and a variety of questions pertaining to the incorporation of portfolio activities into classroom instruction. The questions of impact about which data were collected include the attitudes of teachers and students, perceived effects on achievement, changes in curriculum and instruction, and effects on teachers' evaluations of student performance. The studies of measurement quality include studies of the reliability of scoring (including a planned generalizability-theory analysis

of mathematics) and a validation study focusing on convergent and divergent evidence (using class grades, scores on the state's Uniform Test of mathematics, and standardized test scores).

Data Collection

Data for the Year-2 efforts were obtained through a variety of methods. A statewide sample of principals were interviewed. A sample of classroom teachers in the same schools as those principals were observed and interviewed and provided samples of assignments. Questionnaires were also sent out to a larger sample of teachers and eighth-grade students. The state's contractor will be providing data from the state's uniform test for purposes of validation, and RAND has collected other test scores and grades from the sample of schools in which interviews were conducted. The state agreed to have samples of mathematics portfolios rescored (at the original scoring sessions); RAND provided the design for the rescoring and paid for the required labor. Data for analyzing the reliability of the scoring of the writing portfolios arises naturally from the moderation of teachers' ratings; those data were provided to RAND by the state's testing contractor. Finally, a sample of portfolios were photocopied to explore the possibility of other, content-oriented validation work at a later date.

Samples of Schools

The Year-2 data reported below reflect two samples of schools. The *primary RAND sample* comprises 77 schools selected by RAND to be representative of the state's population of schools that included either fourth or eighth grades. It includes 38 fourth-grade classes and 39 eighth-grade classes, each in a different school. The primary RAND sample was used for the interviews of principals reported below as well as for classroom observations, teacher interviews, and the collection of assignment logs, all of which will be reported in a later publication. Most schools in this sample also provided grades and scores on standardized tests for use in validation studies. The *mathematics Uniform Test sample*, which, as its name implies, was designed for the administration of the state's Uniform Test (UT) of mathematics, was an augmentation of the primary RAND sample, based on the same sample frame and stratified random ordering of schools. It was created because the primary RAND sample would have provided too few

students for reliable estimates of performance on the UT. The uniform test sample is the basis for the Year-2 teacher questionnaire results reported below; the questionnaires were administered with the UT. The uniform test sample will also play an important role in the validation work to be reported in later publications.

For readers interested in the details of the sampling, a brief description is provided in Appendix A.

Results of the Second-Year Principal Interviews

As part of our second-year studies, all principals in our primary sample of schools (see above) were interviewed. Principals were interviewed over the phone because of their geographic dispersion. The response rate among the 77 principals in our final sample of schools was nearly perfect. (However, some schools refused to participate in the study as a whole; those refusals, and the steps taken in response, are described in Appendix A.) One principal asked that we interview the assistant principal instead, and the assistant principal's responses were included in the results presented below. A second principal deferred to others for three of our questions; because the interviewer was eventually directed for that information to the teacher we had observed and interviewed independently, we did not obtain information pertaining to those three questions from that school. All other interviews were complete.

The interview protocols were highly structured. Questions were presented in identical form to all respondents. (See Appendix B for the interview protocol and the instructions for interviewing.) Interviewers were permitted to follow up with *ad hoc* prompts if principals' questions seemed inadequate.

Because of the structure of our interviews, principals' responses fall into two different categories that must be interpreted differently. The answers to some questions can be tabulated as percentages that sum to 100%, but the answers to some others cannot. For example, principals were asked the following question:

> "From your perspective as an administrator, how burdensome, if at all, do you find the portfolio system? If you find it a substantial burden, can you briefly explain what aspects of the program are most difficult for you?"

The answers to the basic question—in this case, whether principals perceive the program as burdensome—can be reported as simple percentages that sum to 100%. If 85% of the principals reported the program to be burdensome, one can infer that 15% did not respond that way.

However, most of our questions also included open-ended probes, the answers to which must be interpreted differently. In the case of the question noted above, principals who found the program burdensome were asked to explain what aspects of the program they found most difficult, and they were free to enumerate as many or as few aspects as they deemed appropriate. We did not prompt them with questions about specific aspects of the program. For example, we report below that:

> "Thirty percent of all principals interviewed commented on what they perceived as unclear expectations, conflicting or tardy information, or unspecified poor communication from the state."

This does *not* imply that 70% did not encounter these problems; rather, it indicates only that 70% did not mention them. For example, some principals may have encountered these problems but reported only others that were more important to them or more salient in their views of the program. This distinction is kept in the phrasing of the results in the following sections. Note that in this case, we reported that 30% "commented on" these difficulties, rather than saying that 30% "encountered" them.

Percentages based on these open-ended probes are useful nonetheless. First, they indicate the *relative* salience or importance of the answers in the views of the respondents. Second, they can sometimes provide a rough indication of the importance of the factors cited. A low percentage need not indicate that an issue was unimportant, but a sizable percentage suggests that it was. In the example given, for instance, the fact that so many principals pointed to problems in the state's administration of the program when asked about burdens implies that problems of that sort were probably substantial.[6]

[6] The percentages described in this section are unweighted. In a variety of test cases, weighting the data to account for non-proportional sampling rates (see Appendix A) made essentially no difference; weighted and unweighted results were within 2 percentage points. Accordingly, we have used the simpler, unweighted results, so that the percentages and counts in the text match.

Use of Portfolios

The state's portfolio program was an innovation in most of the sampled schools. Only three of the 77 principals interviewed stated that their schools had participated in the state's pilot portfolio program in the 1990-91 school year. In addition to those three, 20 additional principals (about a fourth) stated that their schools had worked with portfolios in the past, but their earlier efforts were not as organized or formal as those they have undertaken because of the state's program. Moreover, although a number of principals noted ways in which portfolios were used relatively informally to gauge students' progress, none indicated that their earlier portfolio efforts had been used for formal assessment purposes. Five of the 23 principals who had used portfolios in the past said that teachers retained material informally to demonstrate students' progress to parents during conferences. Principals in 10 schools—13% of the total—noted that they had teachers who kept folders of students' work in order to show students' progress through the year or to substantiate grading decisions.[7]

In a large number of schools, the portfolio system had already spread beyond the two grades specified by the state (fourth and eighth) or is expected to shortly. This suggests that the program is seen as a successful and worthwhile educational intervention by practitioners, despite the fact (explained below) that most of our respondent principals also perceived the program as burdensome. About half of the principals (36 of the 77) stated that teachers in their school in grades other than fourth and eighth are already currently working with the portfolio system. Nine percent of all principals stated that *all* teachers in their school are already keeping portfolios. A few additional principals indicated that they are encouraging other teachers to use portfolios. Moreover, another 10% of the principals interviewed indicated that they are not yet keeping portfolios outside of the grades in which the state requires it but hope to next year.

[7] In five schools, records of the students' performance were passed on to their next teacher to give them a better idea of the level at which a child is performing, but these records were unlike portfolios; most were in the form of "level sheets" that tracked their progress on a set of essential skills.

Burdens Imposed and Support Needed

We approached the question of the burdens imposed by the program by asking three sets of questions: one about principals' overall perceptions of burden from their vantage point as administrators, one about support provided to participating teachers, and a third pertaining to additional support that teachers requested but that principals could not provide. (See question sets 3, 6, and 7 in Appendix B.)

Principal's perceptions of burden. The great majority of principals (86% of our sample) labeled the portfolio program as burdensome. Moreover, some of the remaining 14% reported burdens imposed by the program—such as time required by the program and problems in the state's administration of it—but apparently considered these minor enough not to warrant the term. The principals themselves, however, appear not to bear a large share of the burden they perceive. Despite the wording of our questions, principals more often pointed to burdens borne by teachers than to those they faced themselves.

This perceived burden need not imply a lack of support for the program, however. Twelve percent of the principals labeling the program as burdensome noted (without being asked) that the program was beneficial nonetheless. Perhaps more important is the fact nearly half of our respondents indicated that the portfolio program had already been implemented in their schools outside the grades stipulated by the state, and that yet more intend to expand the program to additional grades shortly. Taken together with comments by a number of principals about positive effects of the program (see below), these findings suggest that the program is perceived by many principals as a worthwhile burden.

The specific difficulties enumerated by principals were diverse. Some complaints were specific to the design or administration of Vermont's program. Others appear to reflect more generic difficulties in using portfolios as the basis of a large-scale assessment. Yet others represent the perhaps inevitable difficulties inherent in implementing any large-scale, systemic reform of educational administration and practice.

Complaints about the state's administration of the program were abundant. Thirty percent of all principals interviewed commented on what they perceived as unclear expectations, conflicting or tardy information, or

unspecified poor communication from the state. (Note that principals were free to list as many or as few difficulties as they believed appropriate and were not asked specifically about the administration of the program or other factors discussed here.) Some criticized the adequacy of training, and a few specifically criticized the training for scoring portfolios. (Other principals, however, labeled the training as good, if time-consuming.) Some of the principals were more specific, noting problems such as unclear guidelines about what materials should be included in the portfolios, confusion over what would be scored, and insufficient teacher training pertaining to scoring criteria. However, comments at this level of specificity were infrequent enough that tabulations would not be meaningful.

A number of comments suggested that the speed with which the program was implemented was a source of stress—a potentially important comment, given the rapid pace of change envisioned in many current proposals around the nation for assessment-based reform. About 8% of the principals explicitly stated that the program was implemented too rapidly. A few of these and an additional 8% commented that the program required difficult rethinking and adaptation by teachers, which is consistent with concern about the pace of change. One in ten principals (12% of those who labeled the program as burdensome) noted that beginning portfolios in two subjects in the same year was burdensome.

A large number of principals complained of burdens not specific to the particular design employed in Vermont—in particular, record-keeping, logistics, and time demands. Many principals did not elaborate about specific time demands, but of those who did, many noted the time required for teacher training. Costs (for substitutes) were mentioned in this context by only a few principals, although when asked about support for teachers, paying for substitutes to permit release time was noted by a large number of them. Perhaps using substitute funds in this fashion was not perceived as problematic by some principals.

Principals' comments about the burdens imposed by the portfolio system in the 1991-92 school year both overstate and understate the likely long-term burdens, albeit to unknown degrees. On the one hand, some of the burdens noted by principals are clearly start-up costs of a new system. The burden of teacher training, for example, should decline over time, and teachers will

gradually become accustomed to teaching and managing their classrooms in ways that facilitate collecting portfolios. Consistent with this, a substantial number of principals—more than a fifth of those who labeled the program burdensome—stated that the burden of the program would likely be at least substantially transient. On the other hand, the portfolio system had been implemented only in two subject areas, and some of the burdens can be expected to grow as the state expands the portfolio system into other subject areas.

Support for participating teachers. The overwhelming majority of principals (71, or 92%) provided special support to their teachers participating in the portfolio project. When asked to specify the types of support provided, the great majority noted release time, but they mentioned other types of support as well. Twelve percent of all principals noted that they provided participating teachers with as much release time as they needed.

Most principals commented that they had provided release time for out-of-school activities. About three-fourths of the principals specified that they provided release time for teachers to attend state-sponsored training sessions. Principals also mentioned providing release time for a wide variety of other activities, but none was mentioned by a majority of the respondents. Thirteen percent of the principals mentioned giving their teachers time to attend state meetings, network meetings and the scoring sessions. Thirteen percent noted providing time for the teachers to spend time discussing the portfolio project with other teachers in their school, district, or supervisory union, in part to enable them to increase the number and types of problems they were including in their portfolios.[8]

Some principals also provided release time for within-school activities. More than a fourth (29%) of the principals noted that they gave their teachers the time to work on their portfolios outside of class. This time was used for things like compiling their students' best pieces, or finding more problems to use with the portfolio program. Some principals stated that it would have been impossible for the teachers to keep up with the extra paperwork without this extra time. A few (5%) mentioned freeing up teachers to conference with the students outside of class about their portfolio work.

[8] Supervisory unions are jurisdictions comprising a small number of local districts.

Within-school support was not limited to release time. Some principals provided support in the form of in-school training sessions. Eight percent reported arranging for trainers to hold additional training sessions in their schools, and 6% reported holding training sessions for other teachers in their area as well. A small number (6%) mentioned arranging extra assistance in the fourth-grade classroom (using either a Chapter 1 aide or other aides).

A few (4%) of the principals reported providing extra time or other assistance so that teachers could develop plans to implement the program in grades other than those (fourth and eighth) stipulated by the state program. This assistance included providing release time for fourth- through tenth-grade teachers to attend the state training sessions and having a Chapter 1 aide train teachers in the third through seventh grades on how to work with portfolios.

Most of the release time was provided by having substitutes cover the teachers' classes. Principals also found other creative ways to provide coverage, by allowing teachers to spend time working on portfolios in lieu of other school duties, or by providing time off as a trade for spending extra personal time working on the portfolios. A few schools also were able to find funds to provide for mileage to attend the training sessions or for tuition in order to get college credit for the summer training course.

The relatively few principals (8%) who reported providing no additional support fall into two categories. Most stated that they provided no assistance only because their teachers did not ask for it. A few, however, said they would not have been able to provide support. Two of the principals interviewed, both in small schools, also teach mathematics. Both stated that in the light of their joint teaching and administrative responsibilities, it was impossible for them to take time off for portfolio related work and therefore they were unable to give themselves the additional time that they may have needed.

We asked if teachers had requested additional support that principals had been unable to provide; only 13% of the principals responded that they were not able to provide some additional type of support the teachers had requested. Three principals reported being unable to provide time requested by teachers, and another three were unable to assist teachers requesting clearer or additional information from the state about what should be included in

portfolios. (A few reported being unable to get in touch with the state-sponsored networks.) A few requests (e.g., one request for an aide) could not be honored for want of funds.

Perceived Attitudes of Teachers, Students, and Parents

We asked principals about the attitudes of teachers, students, and parents toward the portfolio program. All provided substantial comments about teachers' reactions, but some lacked knowledge of students' attitudes, and fewer yet had substantial information about the attitudes of parents.

Attitudes of participating teachers. Principals were asked to note how positive or negative their teachers' attitudes were "on balance," and they were asked to specify the aspects of the program that evoked the most positive and negative reactions.

The great majority (over 70%) of principals characterized the attitudes of their teachers to the portfolio program as mixed. Twenty-three percent mentioned positive feelings on the part of teachers. Only 4% of the principals characterized their teachers simply as negative about the program, and a single principal felt the teachers involved in his school remained neutral.

The aspects of the program that principals said provoked positive responses were myriad. About a fifth of the principals (18%) made generic comments about positive responses to the concept of the portfolio program. A large number (35%) mentioned the measurement value of the portfolios, although their comments often pointed to within-class assessment rather than to comparisons across schools. (A few noted explicitly that teachers' positive attitudes did not extend to the comparative use of the results.) For example, 10% said that their teachers are positive about being more able to see students' progress, and 8% referred to having a better feeling for what children are thinking and doing.

An even larger proportion of principals (47%) mentioned the value of the program as an educational intervention in explaining the positive responses of their teachers. These comments were particularly diverse. Among the many aspects noted were unspecified positive changes in curriculum, better communication and collaboration among teachers, higher levels of thinking and work, students' responsibility for and ownership of the portfolio work, a

broadening of individuals' views of mathematics and of mathematics activities, a movement away from traditional mathematics (by, among others, teachers who otherwise would not have made those changes), and a lessening of "math phobia." Four principals noted state-sponsored training (either the quality of the training or the ready availability of assistance from trainers). A single principal made reference to equity—specifically, the similarity of the program in rural and urban areas.

The aspects of the program that principals said sparked negative responses by teachers were also varied but were more easily categorized. A substantial number of principals (34%) mentioned that the teachers felt the state administered the program poorly. Eighteen percent of all principals mentioned that teachers had problems with communication by the state; specifically, 12% of all principals stated that the communications were poor and unclear, and the other 6% mentioned that information was often late. Others referred to inconsistency in the information provided by the state or to an inability to obtain answers to questions. Eight percent of the principals mentioned that the teachers felt the state's expectations for performance with the portfolios were unclear and that teachers had insufficient guidance.

Some principals referred to difficulties that arose because of the novelty of the program. Ten percent mentioned that the state's time frame for implementation of this program was too short and that full statewide implementation should have been postponed until 1992-93 to give teachers more time to become acquainted with the portfolio program. (Bear in mind that few of our sample schools participated in the 1990-91 pilot program.) A few (6%) said that their teachers were negative because they saw the program as yet another fad that will, like others, disappear.

Many principals mentioned negative attitudes stemming from burdens imposed by the program. Twenty-six percent of principals mentioned that teachers found the time they had to spend on the portfolios to be burdensome. Ten percent mentioned that teachers felt they had to be out of class too often in order to attend the necessary training sessions. Ten percent of the principals mentioned that teachers felt it was inappropriate to have the fourth-grade teachers burdened by both the math and writing portfolio programs. Problems with logistics were also mentioned. Other principals made comments about

teachers feeling frustrated, stressed, or overwhelmed, but without specifying what aspects of the program were perceived as burdensome.

Twenty-one percent mentioned problems the teachers had with the scoring of the portfolios. Aspects of scoring that were cited as provoking negative responses included subjectivity (8%), inconsistency (5%), and difficulties with the training sessions for scoring (6%). (One principal stated that teachers were scared away from assisting in the final scoring because of the overwhelming amount of information covered at the training sessions.)

Attitudes of participating students. A sizable number of principals (about 30%) were either unable to characterize students' responses to the program or stated that students had no particular reaction to the program. Among the 70% who described student responses, however, positive responses far outweighed negative. About 35% of all the principals characterized the students as having mixed feelings about the portfolios, and about a third of them characterized their students' opinions about the portfolio program as positive. Only two principals (3%) stated that their students held predominantly negative feelings about the portfolio program.

About a fifth of the principals noted specific elements of the general portfolio approach about which students felt positive. These elements were diverse. Ten percent of the principals stated that their students are positive about the program because of feelings of pride in or ownership of their portfolio work. Eight percent of the principals mentioned that the students liked the types of problems they were working on; they enjoyed the problem-solving activities and liked finding creative solutions to the answers. Other elements mentioned as sparking positive reactions included doing hands-on projects, using manipulatives, selecting the best pieces, and using language in math activities. A smaller number of principals mentioned students feeling positively about the opportunity for interdisciplinary work, the ability to see their progress over time, a perception that they are thinking at a higher level than before, and the ability to use concepts learned in their portfolio work to solve difficult work in textbooks.

One neutral comment was made by 13 (17%) of the principals. These maintained that students in their schools saw the portfolios not as a distinct

part of the program, but merely as one more thing they have to do during the day.

Principals' negative comments were yet more diverse and more difficult to summarize. Eight percent of the principals mentioned that students do not like the portfolios because of the writing involved; in particular, they reported, students found it difficult to write down their thought processes and disliked doing so. The other student problems mentioned were each mentioned by only one principal; these included having excessive teacher absences because of portfolio training, thinking that other topics were being skipped in order to work with the portfolios, being frustrated that there was more than one answer to many problems, and having anxiety over how the portfolios factored into their grades and how they were going to be scored.

A number of principals noted that students' responses were shaped by the newness of the program. Eight percent noted that they believe students like the portfolio work because of the fact it was new and different, and one principal explicitly stated that only if the students were still excited about the portfolios in three or four years will anyone be able to decide that it was more than the newness of the program that caused the excitement. Conversely, some principals noted that students were initially resistant or negative but became positive over the course of the past year.

Not surprisingly, a number of principals mentioned a belief that the attitudes of students were tied directly to the attitudes of their teachers. Nine percent of the principals made this comment with respect to students' positive attitudes. A few other principals attributed a change from initial resistance to positive attitudes on the part of students to parallel changes in the attitudes of their teachers.

Our interviews with principals were not designed to ascertain whether the attitudes of fourth-grade students differed from those of eighth graders, but the principals' responses offered hints that the attitudes of the younger students many be more positive.[9] Principals in schools with only fourth grades

[9] We sampled a classroom from only one grade in each sample school, but some of our schools included both 4th and 8th grades, and we did not ask principals to contrast the two or to limit their comments to the grade we sampled. We provided a brief questionnaire to eighth-grade students that included questions about their attitudes toward portfolio work, but we did not administer a questionnaire to fourth-grade students because of concerns about insufficient reliability.

were somewhat more likely than those with only eighth grades or both grades to characterize the attitudes of their students as predominantly positive. These patterns reflect very small numbers, however; moreover, fully half of the eighth-grade-only principals characterized their students' attitudes as neutral, which may indicate that principals in secondary schools have less knowledge of students' attitudes. In addition, seven principals explicitly mentioned the two grades in their answers, and these principals generally found the attitudes of fourth graders more positive. Five of the seven labeled the responses of fourth graders as positive, and the other two pegged them as neutral. In contrast, three principals had negative comments about the responses of eighth graders, one reported neutral responses, and three reported positive responses. Moreover, two of the three principals reporting positive responses by eighth graders characterized their initial responses as negative. While one cannot place much confidence in results from only seven schools, these apparent grade differences in the responses of students warrant further exploration.

Parents' attitudes. The majority of our sample of principals had little or no information about parents' opinions of the portfolio program, but most who had information reported that parents are generally positive. Nearly two-thirds (49 principals, or 64%) of the principals stated either that they did not know what parents felt or that parents had not expressed clear opinions to them. Some of these 49 principals reported that they had received no complaints about the program, but none noted having any clear idea if the parents are in favor of the program or not. More than a quarter (29%), however, reported that the parents in their school were positive about the portfolio program. The remaining 8% of principals reported more mixed responses from parents.[10] No principals reported having completely negative reactions from the parents in their schools.

The lack of information about parents' views of the program was a pervasive theme, even among many of the principals who offered an opinion about them. Even many of the principals who said that parents are generally positive about the program noted that they have received only limited feedback from them. This dearth of information may indicate that principals are not the first to receive parental comments on new educational programs, but some

[10] Percentages do not add to 100 percent because of rounding.

principals suggested that it reflected a lack of knowledge or interest on the part of parents. For example, one principal held an information session that was attended by only seven people.

Principals' characterizations of parents' opinions were too diverse to summarize neatly. Ten percent of the principals noted that parents liked elements of the portfolio concept; among them, they mentioned specifically the ability to see their children's work and progress rather than just their test scores or grades; the new way mathematics is being taught; the portfolio systems' record keeping capabilities; and the greater opportunity for parental involvement at home in their children's work. Three principals mentioned that at least some parents feel positively about using portfolios as an assessment tool, in some cases viewing them as a desirable replacement for standardized tests.

The fewer negative parental views volunteered by principals sometimes mirrored the positive. For example, while some principals said that parents welcomed the opportunity for greater involvement in their children's work, a few mentioned receiving critical comments in this regard. Two principals noted that parents had asked that their children not be given the new problem-solving tasks as homework. (One said they view it as "a new torture for parents.") These criticisms may be more important than their small number suggests, because the principals noted that some parents find themselves unable to assist their children with the new types of problems, which raises important questions of equity.

Four percent of the principals reported either that parents see the program as a fad or that they believe it detracts from traditional mathematics and "the basics," or both. Only a single principal mentioned negative parental views about relying on portfolios as a primary means of assessment. This principal believed that the parents were positive about the portfolio concept in general but did not want standardized tests replaced as long as other multiple-choice tests, such as the SAT, will be important in their children's futures.

Perceived Effects on Instruction

Principals were asked whether the program has had either beneficial or detrimental effects on mathematics instruction in their schools. Both

possibilities were explicitly mentioned, and principals were asked to elaborate if they perceived either beneficial or detrimental effects.

A sizable proportion of our principals either had no view of the program's impact on mathematics instruction in their schools or offered none, but the great majority of those who stated a view believe that its effects have been positive. The majority (60%) of all principals in the sample stated that the program has had beneficial effects on mathematics instruction, while 3% perceived mixed effects, and a single principal stated that the program's effects on instruction were basically negative. Another 8% of the principals felt that the program has had no effect on instruction. A quarter of the principals, however, stated that it was too early to accurately assess the impact the portfolio program has had on mathematics instruction in their schools, and 4% would not comment about any effects on instruction.

The positive effects mentioned by principals included a variety of changes in curriculum and instructional methods or styles. Quite a number of principals gave very general answers, such as suggesting that the portfolio system stimulated math instruction, "fired up" tired teachers, encouraged teachers to change outdated methods, or helped to induce desirable curriculum reform. Many were more specific, however, and the effects they mentioned were diverse. One relatively frequent comment (made by 16% of the principals) was that teachers increased their emphasis on problem solving and "flexible thinking." Other principals mentioned specific changes in instructional methods or styles, including a lessened reliance on textbooks, less emphasis on drill and practice, an increased reliance on hands-on learning, increased use of interdisciplinary projects, and an increased emphasis on communication of mathematics. A small number of principals mentioned an increase in group work. Several made comments suggesting that the portfolio system had caused teachers to focus more on students' understanding of and thinking about mathematics.

Although only three principals characterized the instructional effects of the portfolio program as negative, their comments are noteworthy because they are consistent with other comments that were made by a sizable number of principals in response to other questions. All three of these principals maintained that the program was detrimental to mathematics instruction because of the readjustment and loss of time.

Perceived Effects on Traditional Mathematics Performance

We asked principals to comment on the impact of the portfolio program on students' performance on more traditional mathematics activities, but few principals offered any views on this. About half of the principals said they did not know what effects the program had had on that aspect of performance. The responses of an additional third of the principals were not strictly relevant to the question. (Many of the responses that were not fully relevant were positive comments about *non-traditional* skills that students were gaining as a result of the portfolio program. For example, one responded that the program has broadened the scope of students' mathematics ability.)

Only 19% of the principals offered opinions that were truly germane to the question of impact on traditional mathematics performance. Eight (10%) of all principals stated that the portfolios were having a positive impact, while seven of the principals stated that the portfolios were having no impact in this regard. None offered a negative view.

Even though only a few principals had views about this question, their answers were both consistent and intriguing. A commonly voiced concern about the current wave of mathematics reforms is that traditional aspects of competence in mathematics, such as computational skills, might be eroded as activities consistent with the NCTM standards partially supplant traditional mathematics instruction. Yet all eight of the principals who mentioned an effect of the program on traditional mathematics performance said that students are beginning to transfer approaches and skills from the portfolio program to their traditional mathematics work. For example, 5% of the principals stated that the students are beginning to apply the idea of multiple approaches to traditional math problems. One principal stated that being aware and looking at the process of what they are doing, which students experience in their portfolio work, helps them work with traditional activities as well. Another few principals commented that students are now more likely to go over their work more and recheck what they have done as a result of working with portfolios. One stated that the students do not think twice about rechecking their work and finding an alternative way to solve the problem if they find an error.

A few principals commented about the program's effects on students' attitudes toward traditional mathematics work. One said that students have become more enthusiastic about traditional work in geometry. A second said that his fourth-grade teacher reported that the portfolio program has induced a better response to mathematics from students who are normally unresponsive. A third argued the reverse but phrased his answer as speculation; he maintained that students will not want to return to traditional math activities after doing portfolio work and that students will suffer if teachers skip computation to place more emphasis on reasoning skills.

Results of the Second-Year Teacher Questionnaire

Teacher Questionnaires were sent to all fourth- and eighth-grade mathematics teachers in the Uniform Test sample. The questionnaires were packaged with the Uniform Tests. Teachers were instructed to complete the questionnaires while students were completing the Uniform Test or, if that was not possible, as soon thereafter as possible. Questionnaires and testing materials were returned together to an independent contractor for scoring. (See Appendix C for a copy of the 1991-92 Vermont Teacher Questionnaire.)

The response rate for the questionnaire was high. Completed questionnaires were received from 83% of the 172 sampled teachers. The response rate was higher among fourth-grade teachers (90% of the sample of 124) than among eighth-grade teachers (67% of the sample of 48). Response rates on individual questions ranged from 82% to 100%, with the vast majority of questions being completed by more than 95% of the teachers.

Teacher Characteristics

Teachers who responded to the survey had teaching assignments that fit the traditional image of fourth- and eighth-grade teachers. Fourth-grade teachers in the sample could be characterized as generalists, only a few of whom have responsibility for teaching in another class as well as their own. Eighty-two percent of the fourth-grade teachers taught mathematics to a single class with students of heterogeneous ability (93%) and were not mathematics specialists (95%).

In contrast, the typical eighth-grade respondent (70%) taught mathematics to three or more classes and identified himself or herself as a

mathematics specialist (73%). Almost 40% of the eighth-grade classes were grouped by ability level; these were mostly algebra classes (high ability students) and pre-algebra or general math (average ability students).

A minority of eighth-grade teachers had teaching assignments that differed from this pattern. Ten percent of the eighth-grade teachers taught math to only a single class, and nineteen percent taught math to only two classes. Although we do not have information about school organization, we hypothesize that these teachers were working in self-contained classrooms in K-8 or K-12 schools. Consistent with this speculation is the finding that 80% of the eighth-grade teachers who said they were not math specialists were teaching only one or two mathematics classes, and 89% of the eighth-grade teachers who said they were mathematics specialists were teaching math in three or more classes.

Most teachers in the sample had considerable teaching experience, but there was a great deal of variation in experience at both grade levels. Teaching experience ranged from 1 to 35 years. On average, fourth-grade respondents had 15 years teaching experience, and eighth-grade respondents had 17 years. Three-quarters of the teachers had 10 or more years experience, and one-half had more than 16 years experience. There were very few new teachers. Fewer than 10% of fourth-grade teachers and fewer than 20% of eighth-grade teachers had less than five years teaching experience.

Portfolios were new to the great majority of teachers. About one-fifth (22%) of fourth-grade teachers and 35% of eighth-grade teachers had participated in the mathematics portfolio pilot last year.

Description of Participating Classes

Teachers who taught more than one mathematics class were asked to select the first class of the day that was participating in the portfolio assessment and use this as the focal class for responding to the survey. The results below are based on this set of classes.[11]

[11] This procedure may have introduced some unknown bias into the results for teachers who taught more than one class (primarily eighth-grade teachers). Past experience, however, suggests that differences due to this type of selectivity are minor, and we do not believe they had any systematic effect on the results.

Teachers exercised some selectivity in choosing classes to participate in the portfolio assessment in 1991-92. Most teachers restricted participation in the portfolio assessment to one or two classes. In grade 4, only 5% of the teachers introduced the mathematics portfolios to more than a single class, though 18% taught mathematics in more than one class. One-half of the teachers in grade 8 selected only a single class to participate in the portfolio assessment, and another 30% restricted participation to only two classes. This selectivity was present regardless of teachers' past experience with portfolios.

In grade 4, the vast majority of classes participated in the writing portfolio assessment as well as the mathematics portfolio assessment. Fewer than 5% of fourth-grade teachers opted to have their students participate in the mathematics portfolio assessment but not the writing portfolio assessment. Few of the grade-8 teachers were responsible for teaching writing, and it is impossible to know on the basis of this survey what percent of students participated in one assessment but not the other.

In grade 4, the typical class consisted of 19 students of heterogeneous ability level, all of whom were compiling mathematics portfolios. In grade 8, the typical class had 17 students of mixed ability, all of whom were compiling mathematics portfolios. In all but 18% of fourth-grade classes and 8% of eighth-grade classes, all students participated in the portfolio program. (Teachers were not asked to explain the reasons for the exclusions in the other classes.) Thirteen percent of the eighth-grade classes were identified as algebra classes, 35% as pre-algebra, 41% as general mathematics, and the remaining 11% had other titles.

Time Devoted to Portfolios

Teachers devoted considerable time to the portfolios, both in class and outside of class. On average, teachers devoted six hours per week to the portfolios: two to three hours per week preparing for portfolios, a similar amount of time in classroom portfolio activities, and about one hour per week scoring or evaluating portfolios. Although there was wide variation in the amount of time reported, and such retrospective estimates may not be highly reliable, these results suggest that portfolios typically required a substantial time commitment.

Teachers estimated the amount of time they spent on five distinct elements of portfolios: finding tasks, preparing lessons, conducting lessons, helping students, and scoring or evaluating portfolios. Estimates of time devoted to each of type of activity ranged from zero to more than 20 hours per month (see Table 1). In fact, two teachers reported spending 60 hours per month conducting portfolio lessons.[12]

Table 1

Teacher Time Spent on Portfolio Activities (Hours per month)

ACTIVITY	GRADE 4 Mean (SD)	GRADE 8 Mean (SD)	OVERALL Mean (SD)
Finding appropriate tasks and/or materials	6.0 (6.7)	5.9 (5.2)	5.9 (6.3)
Preparing portfolio lessons	6.3 (5.1)	6.9 (6.8)	6.5 (5.6)
Conducting portfolio lessons	8.8 (5.5)	10.3 (14.5)	9.2 (8.9)
Helping students organize/ manage their portfolios	3.7 (4.5)	2.7 (2.8)	3.4 (4.1)
Scoring/evaluating the contents of portfolios	4.6 (7.1)	6.7 (7.6)	5.2 (7.3)
Total teacher time	28.9 (22.0)	33.2 (27.2)	30.1 (23.6)

Teachers also were asked to indicate the amount of class time students spent doing portfolio tasks, revising tasks, and organizing portfolios. On average, students spent about 14 hours per month on portfolios: about 7 hours doing portfolio tasks, about 4 hours per month revising them and about 3 hours per month organizing their portfolios (see Table 2). Fourth-grade students spent significantly more time doing and revising tasks than eighth-grade students.

Training and Support

Virtually all (over 95%) of the fourth- and eighth-grade teachers received some type of formal, state-sponsored training in the implementation of the

[12] It is not clear what portion of this variation can be attributed to differences in the accuracy of teachers' answers, as opposed to variations in the implementation of the portfolio program.

Table 2

Classroom Time Spent on Portfolio Activities (Hours per month)

ACTIVITY	GRADE 4 Mean (SD)	GRADE 8 Mean (SD)	OVERALL Mean (SD)
Doing portfolio tasks for the first time	7.8 (5.8)	5.3 (4.7)	7.1 (5.8)
Revising or rewriting portfolio tasks	4.1 (5.7)	2.2 (3.4)	3.6 (5.2)
Organizing/managing portfolios	3.0 (3.3)	2.3 (2.9	2.8 (3.2)
Total Classroom Time	15.0 (13.0)	9.9 (8.9)	13.7 (12.2)

mathematics portfolios, and approximately three-quarters of the teachers reported that the training prepared them adequately well to work with the math portfolios.

The Preparation for Scoring workshops were the most widely attended: the vast majority of teachers (88%) participated in these workshops during 1991-92. In addition, approximately 30% of fourth-grade teachers and 45% of eighth-grade teachers attended the summer or fall Math Institutes, and one-half of the teachers participated in some form of network training activity at their local school. In fact, 64% of fourth-grade teachers and 81% of eighth-grade teachers participated in more than one formal staff development activity. Approximately one-half of the teachers reported that they also received informal support from local teachers, but only 10% of teachers received informal support directly from their school principals.[13]

How well did the available training prepare teachers to implement the mathematics portfolios? Approximately one-quarter of the teachers reported that their training made them "well prepared" or "very well prepared" to work with the mathematics portfolios. Approximately one-half indicated that the training prepared them "adequately," and approximately one-quarter said the training prepared them "poorly" or "very poorly." Eighth-grade teachers felt themselves to be somewhat better prepared than fourth-grade teachers. This

[13] Note that principals most often reported giving support to participating teachers in the form of release time for participation in training offered by others (see the Principal Interview results above).

may be due to their greater knowledge of mathematics, to particular characteristics of the workshops, or to both factors.

In addition to the overall rating, teachers indicated the level of preparation they received from each of the training and support activities they attended. The average response for each of the six types of training was "adequate" or better (see Table 3). The Summer Institute received the highest rating from teachers who attended; almost 90% of the teachers said they were at least adequately prepared to work with the mathematics portfolios after attending the institute; 45% felt they were "well prepared" or "very well prepared." The Preparation for Scoring workshop was the next most highly rated, followed closely by informal principal support, informal local teacher support, formal network training sessions, and the fall Math Institute. In most cases, the responses of eighth-grade teachers and fourth-grade teachers were similar.

Teachers relied on a number of different sources for ideas for portfolio tasks. The *Resource Book* was the most frequently used source of ideas for portfolio assignments.[14] More than two-thirds of the fourth-grade teachers and one-half of the eighth-grade teachers referred to it regularly. The next most common sources of portfolio tasks were textbooks, teachers' own ideas, and ideas from other teachers at the school. Materials shared by the Network leader and handouts from the summer and fall institutes were consulted regularly by only about 20% of the teachers at either grade level.

Changes in Curriculum and Instruction

One of the primary purposes of the portfolio assessment was to influence classroom practice. We asked teachers how much they changed their coverage of specific mathematical topics and their emphasis on selected aspects of problem solving and mathematical communication as a result of the portfolio assessment. Consistent with the goals of the program, teachers at both grade levels reported marked increases in the time devoted to teaching the topics of problem solving and patterns/relationships. More than three-quarters of the teachers spent "somewhat more" or "much more" time teaching problem-solving strategies; and approximately one-half spent somewhat more or much more time teaching patterns/relationships (see Table 4).

[14] Vermont Department of Education, 1991c.

Table 3

Teachers' Ratings of Training and Support Activities (% of Teachers)

TYPE OF ACTIVITY	LEVEL OF PREPARATION		
	Poorly or Very Poorly	Adequately	Well or Very Well
	Grade 4		
Summer Math Institute	7	33	59
Fall Math Institute	32	39	29
Preparation-for-Scoring	13	48	40
Network training	23	54	23
Local teacher support	24	48	28
Principal support	0	55	45
	Grade 8		
Summer Math Institute	21	21	57
Fall Math Institute	13	40	47
Preparation-for-Scoring	9	31	59
Network training	15	31	54
Local teacher support	7	60	33
Principal support	0	100	0

Note. Percentages may not sum to 100 due to rounding errors.

In contrast (and perhaps surprisingly), teachers reported only limited changes in the amount of time devoted to computation, estimation, and measurement/geometry. Over one-half of the teachers devoted the same amount of time to these topics in 1991-92 as in the last year prior to the portfolios. Among the remaining teachers, the trend was to devote somewhat less time to computation and somewhat more time to estimation and measurement. Only minor differences were observed between fourth- and eighth-grade teachers in the time they spent on these mathematical topics.

Table 4

Change in Class Time Devoted to Mathematical Topics (% of Teachers)

TOPIC	TIME ON TOPIC		
	Somewhat or Much Less	About the Same	Somewhat or Much More
	Grade 4		
Computation and algorithms	49	53	7
Estimation	8	55	37
Patterns/Relationships	5	46	48
Measurement/Geometry	15	57	27
Problem Solving Strategies	2	16	82
	Grade 8		
Computation and algorithms	31	66	3
Estimation	3	63	33
Patterns/Relationships	10	40	50
Measurement/Geometry	3	73	23
Problem Solving Strategies	0	23	77

Note. Percentages may not sum to 100 because of rounding errors

When we focused more narrowly on communication and problem-solving activities, changes in teachers' emphases were even more striking. Overall, there was an increase in emphasis on all four types of communication activities: describing personal experiences relating to mathematics; having oral discussion of mathematical topics; making charts, graphs and diagrams; and having students write reports about mathematics. These changes were more pronounced for fourth-grade teachers than for eighth-grade teachers.

Written reports of students' mathematics-related projects showed the most dramatic change, receiving either "somewhat more" or "much more" emphasis from 66% of the fourth-grade teachers and 75% of the eighth-grade teachers. Making charts, graphs and diagrams and holding oral discussions also received somewhat more emphasis than in the past from most teachers. One-half of the fourth-grade teachers also placed at least somewhat more

emphasis on describing personal experiences relating to mathematics, but this was not the case for eighth-grade teachers, who reported changing little in this regard.

As one might expect, teachers also devoted more emphasis to most aspects of problem solving. Only single-step word problems did not receive increased emphasis from the vast majority of teachers. (Two-thirds of the teachers gave single-step word problems the same emphasis as previously.) In contrast, three-quarters of the teachers placed greater emphasis on applying math knowledge in new situations and to logic or reasoning problems, and one-half gave more emphasis to multi-step word problems and to collecting, analyzing and reporting data. These results are consistent with the goals and emphases of the new assessment program.

During training, teachers had an opportunity to review examples of three types of tasks they might assign to students to generate work suitable for inclusion in the portfolios: puzzles, applications and investigations. (See the discussion of the assessment program above for a brief description of these three categories of tasks.) Applications were the type of task teachers assigned most frequently in implementing the portfolio program. Over 60% assigned application problems at least once per week. By comparison, only 46% assigned puzzles at least once a week, and only 34% assigned investigations that frequently. Puzzles and investigations were most commonly assigned between one and three times per month.

In addition to curricular emphases, we also asked teachers about changes in instructional practices. Specifically, teachers were asked whether they changed the pattern of student grouping they used for mathematics. About one-half the teachers said more work was done in small groups and in pairs than in prior years. In comparison, about 70% of the teachers used whole group instruction and individual work about the same amount as they had in the past. A greater proportion of eighth-grade teachers than fourth-grade teachers reported changes in the way they grouped students for instruction.[15]

[15] Again it is difficult to interpret these grade level differences. One plausible explanation is that fourth-grade teachers already used a variety of instructional groupings prior to the portfolios, so fewer changes were made, whereas eighth-grade teachers relied on whole class instruction almost exclusively.

Preparation and Revision of Portfolio Entries

Because of the flexible nature of the mathematics portfolio assessment and the lack of strict guidelines for the composition of portfolios and the preparation of individual pieces, each teacher exercised considerable discretion over the work produced by his or her students. One result of this freedom was that mathematics portfolios were not implemented in the same way by all teachers. For example, teachers established different rules for revision and collaboration, and they emphasized different aspects of quality in the final work products.

Teachers tried to emphasize students' interests in selecting activities for the portfolios. Almost 90% of the teachers placed "moderate" or "heavy" emphasis on having work be important or interesting to the students. However, interest was not promoted at the expense of correctness. Eighty percent of teachers placed as much emphasis on mathematical correctness as on student interest. Teachers also used examples from the Resource Guide as models of what student work products should look like.

Many teachers set ground rules for the preparation of portfolio pieces. Sixty-five percent of the fourth-grade teachers and 44% of the eighth-grade teachers placed limits on the amount and/or type of assistance students could received from parents or other adults. Almost all teachers who set limits on adult assistance required students to acknowledge the assistance they received. This was true for about two-thirds of the fourth-grade teachers. (In both grades one-half of the teachers who set limits insisted that students describe the nature of the assistance they received while the other half required only that students indicate some help was given.)

We do not know if specific ground rules were established for revising student work before putting it into the portfolios. However, teachers reported that student work was revised once, on average, before being placed in their portfolios. There were small differences between teachers in this respect: approximately one-quarter of the teachers said student portfolio pieces were not revised at all, approximately one-half said they were revised once, about 20% said they were revised twice, and fewer than 10% said portfolio work was revised three times or more. This result may allay some fears that portfolio pieces were overly rehearsed.

Teachers reported that students played a key role in the selection of best pieces. Students and teachers collaborated in the selection of best pieces in two-thirds of the classes, and student alone were responsible for selection in the remainder. These patterns were generally the same for both grade levels.

Neatness was the dimension on which fourth- and eighth-grade teachers differed most. Fourth-grade teachers placed somewhat less importance on neatness of appearance than on interest, correctness, and matching models from the *Resource Book*. Eighth-grade teachers placed as much emphasis on neatness as on interest and correctness, but placed less emphasis than fourth-grade teachers on resembling examples from the *Resource Book*.

Teachers' Evaluations of Student Performance

More than 80% of the teachers indicated that they had changed their opinion of students' mathematical abilities on the basis of students' portfolio work. The majority of teachers reported changing their opinions of students' abilities only a small amount, but more than one-third of the teachers reported that they changed their opinions a "moderate amount," and nearly 10% changed their opinions "a great deal." The results were similar for fourth- and eighth-grade teachers. Slightly greater change was reported among teachers who participated in the portfolio pilot in 1990-91 than among those who used portfolios for the first time in 1991-92. We cannot discern, however, whether these results reflect changes in students' performance (because of the different type of tasks they were given), changes in teachers' perceptions (because of the new information provided to them by portfolios), or both.

Many teachers had difficulty generalizing about differences in students' performance between portfolio tasks and regular mathematics work. One-quarter of the teachers said performance varied too greatly across tasks or students to make overall comparisons between portfolios and regular math work. Over one-third of teachers reported that students' performance was about the same on the two types of math assignments. Of the remaining one-third, most reported that students did worse on portfolio tasks than regular mathematics work. In grade 4, 27% of the teachers said students performed worse on portfolio tasks compared to 13% who said students performed better. In grade 8, 24% of teachers reported worse performance on portfolios while only 3% reported better performance on portfolios.

Effects of Portfolios

Teachers reported that the mathematics portfolios had a number of positive effects on themselves and their students. For example, more than one-half of the teachers said they were frequently more enthusiastic about teaching math, and over 90% were more enthusiastic at least occasionally (see Table 5).

Table 5

Frequency of Positive Effects (% of Teachers Reporting)

ISSUE	FREQUENCY OF POSITIVE EFFECTS		
	Rarely or Never	Occasionally	Often or Always
Grade 4			
I am more enthusiastic about teaching math	15	29	56
Goals of math instruction are improved	10	33	57
Math is more closely linked to other subjects	14	41	45
Students' attitudes toward math improve	19	38	43
Students are learning more mathematics	14	35	51
Low ability students are more successful	28	40	33
Grade 8			
I am more enthusiastic about teaching math	11	38	51
Goals of math instruction are improved	14	57	30
Math is more closely linked to other subjects	11	46	43
Students' attitudes toward math improve	32	41	27
Students are learning more mathematics	14	51	35
Low ability students are more successful	38	46	16

Note. Totals may not sum to 100% due to rounding errors.

Similarly, over 40% said the following positive effects occurred frequently:

- The goals of mathematics instruction are improved;

- Math is more closely linked to other subjects;
- Students' attitudes towards math improve; and
- Students are learning more mathematics.

Teachers who participated in the 1990-1991 portfolio pilot were only slightly less enthusiastic this year than other teachers.

Teachers perceived portfolios as having both positive and negative effects on the performance of low ability students. Approximately one-quarter of the teachers reported that low ability students were frequently more successful as a result of the portfolios. Low ability students experienced greater success "occasionally" in another 40% of the classrooms (see Table 5). However, on a separate question, teachers indicated that low ability students often experienced difficulty with portfolio tasks (see Table 6). Virtually all fourth-grade teachers and 80% of the eighth-grade teachers indicated that low ability students had difficulty with portfolio tasks at least occasionally.

More fourth- than eighth-grade teachers reported instructional benefits from the portfolio program. For example, over one-half of fourth-grade teachers indicated that the goals of math instruction were improved often or always compared to one-third of eighth-grade teachers. Similarly, one-half of the fourth-grade teachers reported that students often learned more mathematics as a result of the portfolios compared to one-third of the eighth-grade teachers. One-half of the fourth-grade teachers said attitudes had improved often or always compared to one-quarter of the eighth-grade teachers. Differences between fourth- and eighth-grade teachers' responses to the other two items (teachers' enthusiasm for mathematics and links to other subjects) were minor.

Portfolio-Related Problems

In the view of teachers, the greatest problems created by the portfolios were not what to do, but when to do it. Finding time to cover the full mathematics curriculum was the most widespread problem, followed closely by finding time to prepare portfolio lessons. Over 80% of fourth-grade teachers and over 60% of eighth-grade teachers often had difficulty covering the required curriculum. Sixty percent of both groups often lacked the time to prepare portfolio lessons (see Table 6).

Table 6

Frequency of Portfolio-Related Problems (% of Teachers Reporting)

ISSUE	FREQUENCY OF PROBLEM		
	Rarely or Never	Occasionally	Often or Always
Grade 4			
I don't understand what I'm expected to do	24	50	26
I don't have enough training in how to do it	25	39	36
I have difficulty finding appropriate tasks	25	42	33
I lack time to prepare portfolio lessons	15	25	59
Not enough time to cover the full math curric.	4	15	81
Low ability students have difficulty with tasks	3	37	60
Students don't understand what to do with tasks	9	45	45
Students don't know how to solve problems	7	43	50
Students not interested in portfolio tasks	29	50	21
Grade 8			
I don't understand what I'm expected to do	27	57	16
I don't have enough training in how to do it	38	43	19
I have difficulty finding appropriate tasks	41	41	19
I lack time to prepare portfolio lessons	14	22	65
Not enough time to cover the full math curric.	11	24	65
Low ability students have difficulty with tasks	19	31	50
Students don't understand what to do with tasks	11	50	39
Students don't know how to solve problems	8	57	35
Students not interested in portfolio tasks	16	51	32

Note. Totals may not sum to 100% due to rounding errors.

In addition, most teachers reported occasional difficulty knowing what they were expected to do or how to do it. This suggests that additional information and training are needed. Three-quarters of the fourth-grade teachers and two-thirds of the eighth-grade teachers felt they lacked adequate training at least occasionally. Teachers who participated in the portfolio pilot last year reported the same feelings, though with slightly less frequency.

Teachers perceived students to be interested in the portfolio tasks. However, they reported that students frequently had difficulty solving portfolio problems: one-half of the fourth-grade teachers and over one-third of the eighth-grade teachers said students often did not know how to solve problems.

References

Mills, R.P., & Brewer, W.R. (1988). *Working together to show results: An approach to school accountability in Vermont.* Montpelier: Vermont Department of Education.

Vermont Department of Education. (1990). *Vermont Writing Assessment: The pilot year.* Montpelier: Author.

Vermont Department of Education. (1991a). *Looking beyond "The Answer:" The report of Vermont's Mathematics Portfolio Assessment Program.* Montpelier: Author.

Vermont Department of Education. (1991b). *"This is my best:" Vermont's Writing Assessment Program, pilot year 1990-91.* Montpelier: Author.

Vermont Department of Education. (1991c). *Vermont Mathematics Portfolio Project: Resource book.* Montpelier: Author.

Vermont Department of Education. (1991d). *Vermont Mathematics Portfolio Project: Teacher's guide.* Montpelier: Author.

Vermont Department of Education. (nd). *"This is my best:" Vermont's Writing Assessment Program, Pilot year 1990-91.* Montpelier: Author.

APPENDIX A: DESCRIPTION OF THE STUDY SAMPLES

APPENDIX A: SAMPLES

This appendix provides a brief description of the two samples that provided the basis for most of the CRESST/RAND study during the 1991-92 statewide implementation of the Vermont assessment program.

The primary RAND sample. This comprised separate samples of 40 eighth-grade and 40 fourth-grade classes, each of which was designed to be representative of the state's population of mathematics classes in that grade (with the exceptions, explained below, of some deliberate over-representation of small and high-poverty schools). We sampled on a school basis, and a single mathematics class (for either the fourth or eighth grade) was selected from within each sample school to be the focus of all class-based data.

To create a sampling frame, we divided the 246 fourth-grade classes and 125 eighth-grade classes in Vermont separately into 6 different cells based on school size (specifically, within-grade enrollment counts) and poverty rates. (The proxy for poverty rates was an adjusted proportion of students in each township receiving AFDC.) Vermont has many small schools, so the strata for school size were 1-14, 15-24, and 25 and more students. Poverty strata were less than 13% and greater than or equal to 13%. This break was determined on the basis of frequencies and subjective judgments of the poverty of selected townships. School size was used as a stratification variable to ensure a sufficient sample of very small schools, reflecting a concern that the portfolio program might be more difficult to implement in them. Poverty was used as the second stratification variable because of the state's concern about the impact of the program on equity.

All schools in each cell were randomly ordered for both initial selection and selection as replacements for refusals. We first drew an eighth-grade sample; we then sampled a second time for fourth grade, using different frequencies because of the different distribution of fourth grades across our stratification variables and excluding all schools that had been selected for the eighth-grade sample. Vermont has a sizable number of K-8 schools, and 9 schools that would have been selected for the fourth-grade sample were

dropped because they had been selected for the eighth-grade sample. Three other schools were also dropped before we contacted them.[16]

An additional 23 schools were dropped for various reasons after we contacted them to ask for their participation. All but three were replaced (those refused too late to be replaced). We considered 9 of these 23 relatively unproblematic in terms of their likely effects on the quality of the sample. Four schools, including 2 in Burlington, were replaced because they chose not to participate in the portfolio program this year. Two schools were replaced because the selected teachers were absent on or had just returned from maternity leave. The other 3 schools were dropped because of teachers' personal problems.

We considered the non-participation of the remaining 14 selected schools more problematic. Eleven teachers or principals simply refused to participate in our study. One one-room school refused to participate because the teacher felt that she did not have the resources available to participate. The other 2 dropped schools were in the same district; the teachers were initially uncooperative, and the superintendent, noting that the district has many one-room school houses with "oodles of problems," declined to intervene.

Final sample counts and the proportions of schools sampled from within each cell were very similar to those in the original sampling design. Our sample of 38 fourth-grade schools represents 15% of the total number of Vermont schools that include a fourth grade. Most of our schools had fourth-grade enrollments of less than 25, but in that respect the sample mirrors the state as a whole. To include a reasonable number of high-poverty schools, however, we sampled them at approximately twice the rate of low-poverty schools, regardless of school size (Table A1). Our sample of 39 eighth-grade schools represents nearly a third (31%) of such schools statewide. To ensure a reasonable distribution of schools across our strata in grade 8 required substantially non-uniform sampling rates across cells (Table A2). Very small and small high-poverty schools were substantially over-represented in our sample, while low-poverty large schools were substantially underrepresented.

[16] One school was dropped because its location would have made observation very difficult. Two other potential replacement schools were dropped because they were one-room schoolhouses with extremely small within-grade enrollments. Our initial contacts indicated that gaining participation by the smallest one-room schools was difficult, and we excluded schools with fewer than six students in the target grade from our replacement list.

Table A1

Grade 4 Final Primary RAND Sample

Within-grade Enrollment		Poverty Rate <13%	Poverty Rate >=13%
Very small (<=14)	Number of schools sampled	7	5
	Percent of schools sampled	14%	26%
Small (15-24)	Number of schools sampled	4	4
	Percent of schools sampled	11%	22%
Large (>=25)	Number of schools sampled	8	10
	Percent of schools sampled	11%	21%

Table A2

Grade 8 Final Primary RAND Sample

Within-grade Enrollment		Poverty Rate <13%	Poverty Rate >=13%
Very small (<=14)	Number of schools sampled	3	5
	Percent of schools sampled	30%	56%
Small (15-24)	Number of schools sampled	6	5
	Percent of schools sampled	43%	50%
Large (>=25)	Number of schools sampled	10	10
	Percent of schools sampled	20%	36%

Our participation rates within the final sample were excellent. We were able to conduct principal interviews with all 77 of the schools remaining in our sample, and we were able to conduct observations and teacher interviews in 75.[17]

[17] Two of the fourth-grade observations had to be canceled because logistical problems precluded observing until after their portfolios had already been scored in May. We believed that observations held after scoring would threaten the integrity of the findings more than the loss of the two classes.

The Uniform Test sample. The primary RAND sample included too few students for reliable estimates of student performance on the Uniform Test (UT) of mathematics. Accordingly, Advanced Systems, the contractor for the UT, augmented the primary RAND sample for purposes of the UT, a short student questionnaire that accompanied the test in grade 8, and the teacher questionnaire reported below. The augmentation was designed to obtain a minimum of 35% of the students in each of the six strata in the RAND sampling frame. All students in fourth and eighth grade in each school were tested. Where the RAND sample provided more than 35% of the students in the stratum, it was left intact; where it failed to provide 35%, it was augmented using the random number sequence used in creating the primary RAND sample.

APPENDIX B: YEAR-TWO PRINCIPAL INTERVIEW AND INSTRUCTIONS

PRINCIPAL INTERVIEW GUIDE
RAND/CRESST Study of the Vermont Assessment Program

INSTRUCTIONS

Please familiarize yourself with this form before beginning interviews.

The person interviewed should always be the principal, *even if that means a call back*, unless there is an overriding reason. If there is an overriding reason the principal cannot be interviewed, such as a protracted absence, the interview should be conducted with the person who has administrative responsibility for the school. Note that in many small schools, the principal may also be a classroom teacher and may by chance be the sampled teacher.

You may append additional sheets if needed to put down the results of the interview.

If questions arise concerning their participation in the study or the State Department of Education's concerns, the interviewee can call Ross Brewer, Director of Policy and Planning, (802) 828-3135. If questions arise about our study that you are unable to answer, you may forward the questions to RAND by calling Ed Deibert or Dan Koretz at 202-296-5000. If questions arise concerning *operation of the assessment program* (e.g., which schools are being tested or how results will be reported), refer them to Ross Brewer.

If at some point the interviewee asks how these interviews will be used or whether they will get feedback, explain that individual interviews will be held confidential but that summaries of the collected responses will be a part of the reports of the study issued later this year and next.

After the following "INTRODUCTION" is a section entitled "BASIC INFORMATION" (p. 3). Please fill out as much of that section as you can before continuing. Also, please enter the school name at the bottom of every page.

INTRODUCTION

Introduce yourself by name and explain that you are calling for RAND's independent evaluation of the Vermont assessment program and that you would like about 25 minutes of their time to ask them questions about the program. If the present time is not convenient, schedule a specific time to call back.

The principal should have received a letter from Rick Mills, the Commissioner of Education, explaining the study and asking them to participate. If they have not received the letter or request clarification of the study, you can explain the following points:

1. The State Department of Education requested this evaluation of the assessment program. It is being conducted independently by RAND, a non-profit research organization with extensive experience in education research, and is not funded by the state.

2. The study will evaluate the assessment program and its impact. The results of the study will be very important to those working on the program—including many Vermont teachers—as they try to improve it over the coming years.

3. **We will *not* evaluate any individual schools or school staff. We will *not* release any information about individuals or specific schools to the State Department of Education or anyone else. The information participants provide us is strictly confidential.**

4. To minimize burden on participating schools, only one teacher has been selected from each school. Letters of explanation have already gone out to those teachers. We are requesting:

   - Mid-year and final grades.
   - A two-week log of mathematics assignments.
   - An observation of one mathematics lesson, followed by a brief interview.
   - An end-of-year interview (with a subset of selected teachers).
   - Standardized test scores for the observed class.
   - Teacher questionnaires. (These will be given to all teachers administering the state Uniform Test of Mathematics, not just the one teacher we selected. The questionnaires will be packaged with the Uniform Test and can be filled out during the testing session).

***In all cases**, please explain the following points:*

1. The school and teacher were carefully selected to get a representative sample of the state's schools. Their participation is therefore essential if we are to maintain the integrity of the evaluation.

2. We recognize that participation in the study imposes some burdens, and on our own behalf and on the behalf of the Vermont teachers and administrators developing the program, we want to thank them for their willingness to participate.

BASIC INFORMATION

To be filled out in advance to the extent possible. Ask all questions necessary to fill out items that you cannot fill out in advance. Immediately after concluding the interview, add the school name to the footer at the bottom of each of the following pages and at the top of each blank page at the back that you use.

Interviewer name: _____

Date of interview: _____ Grade sampled: _____ (4 or 8)

Name of school: _____

Name of LEA or Supervisory Union (if a union school):

_____

Person interviewed: _____

Position of interviewee, if *not* principal: _____

Is the interviewee both principal *and* a classroom teacher? _____ If so, does he or she teach mathematics regularly in the sampled grade? _____

Teacher initially selected for observation (take from master list):

_____

Phone: (802) _____ - _____ Fax: (802) _____ - _____

QUESTIONS ABOUT STANDARDIZED TESTS

Ask the following questions about standardized achievement test batteries. ***IF the interviewee asks why we are interested in this, explain simply that test scores are one of the many pieces that can be used to evaluate the portfolio system.***

Will your district administer a standardized achievement test to the sampled grade this year? _____

If so, please answer the following questions:

Which test will be administered (name and form, if known):

_____

When will the test be administered? _____

When will the results be back in the district? _____

If so, can we obtain scores from the interviewee? _____

[If questions about release of test scores arise, you can explain that the Superintendent will be receiving a letter from the Commissioner of Education explaining why such data should be made available in this instance, and that the Superintendent and the Department will discuss any difficulties with release of the data.]

If not, whom should we see about scores? (Name, title, and phone number)

_____

SUBSTANTIVE QUESTIONS

Please read these questions ***as written****. Jot your notes about the answers on the additional pages in the back. When the interview is over,* ***immediately*** *transcribe your notes into legible form on those pages and add the name of the school at the top of each page you use. You may ask for clarification at any point if the answers seem incomplete or unclear to you, but allow respondents to keep their answers brief.*

1. Were portfolios or similar records (such as ongoing logs) kept in your school's mathematics classes before the State requested mathematics portfolios? If so, please explain briefly what was done and how frequent or extensively those efforts were.

2. Are mathematics portfolios currently being kept by teachers in your school in grades other than 4 and 8?

3. From your perspective *as an administrator*, how burdensome, if at all, do you find the portfolio system? If you find it a substantial burden, can you briefly explain what aspects of the program are most difficult for you?

4. What is your perception of the attitude of *participating teachers* in your school toward the portfolio system? On balance, how positive or negative have their reactions been, and can you note briefly what aspects of the system have evoked the most positive and negative reactions?

5. What is your perception of the attitude of *participating students* in your school toward the portfolio system? On balance, how positive or negative have their reactions been, and can you note briefly what aspects of the system have evoked the most positive and negative reactions?

6. Do participating teachers in your school receive any sort of special support—instructional, logistical, or other—for portfolio-related work? An example might be limited released time to select and compile "best pieces." If so, can you please briefly note what types of support they receive?

7. Have participating teachers requested additional support that has not been feasible to provide? If so, could you note what they have requested?

8. Has the portfolio system had *beneficial or detrimental* effects on mathematics instruction in your school? If so, can you elaborate briefly?

9. *(if not answered in #8)*: In your opinion, has the portfolio system had any impact on students' performance on more traditional mathematics activities?

10. What do parents think of the portfolio system?

CONCLUSION

Please thank the principal on behalf of both RAND and the Vermont portfolio team for their participation in the study.

APPENDIX C: YEAR-2 TEACHER QUESTIONNAIRE

CARD 01

Vermont Teacher Questionnaire

Dear Teacher,

This questionnaire will help us evaluate the Vermont portfolio assessment program. It asks for information about the implementation of the mathematics portfolio program and its effects on curriculum and instruction.

This questionnaire is anonymous and will not be used to evaluate individual teachers or schools. *Even though your questionnaire will be collected with your students' tests, it will be separated from them, and no record will be kept linking questionnaires to schools or teachers.*

Please complete the questionnaire today and return it with the students' completed Uniform Tests. If possible, please complete this questionnaire while your students are working on the Uniform Tests. If that is not possible, please complete it later in the day so that it can be returned to ASME with the Uniform Test books and other assessment materials.

Thank you.

Experience

1. How many years of teaching experience have you had?

YEARS ☐☐ 4-5/

2. In which grade(s) do you currently teach mathematics?

(Check all that apply)

1	2	3	4	5	6	7	8	9	10	11	12
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

6-17/

3. In how many classes do you teach mathematics?

(Check One)

1	2	3	4	5	6
☐	☐	☐	☐	☐	☐

18/

4. Do you specialize in teaching mathematics?

(Circle One)

Yes 1 19/
No 0

5. Did you participate in the mathematics portfolio pilot last year?

(Circle One)

Yes 1 20/
No 0

6. How many of your classes are participating in the mathematics portfolio assessment this year?

(Check One)

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

21/

7. How many of your classes are participating in the writing portfolio assessment this year?

(Check One)

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐

22/

8. Apart from the demands of the portfolios, how prepared do you feel to teach the subject matter covered in your mathematics class(es)?

(Circle One)

Very well prepared	5	23/
Well prepared	4	
Adequately prepared	3	
Somewhat unprepared	2	
Very unprepared	1	

PLEASE FOLLOW THESE GUIDELINES TO SELECT ONE CLASS TO USE TO ANSWER THE REMAINDER OF THE QUESTIONS:

If only one class is participating in the mathematics portfolio assessment, consider it to be the ***designated class*** for this survey and answer the following questions for that class. If MORE THAN ONE of your classes is participating in the mathematics portfolio assessment, please use the FIRST PARTICIPATING CLASS OF THE DAY as the **designated class** for this survey.

9. Which grade levels are included in the designated class?

(Check all that apply)

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐

24-35/

CARD 01

10. Is the designated class grouped by ability (i.e., the class includes only students within a particular range of abilities)?

(Circle One)

Yes 1 *36/*
No 0

IF YES, how would you describe the ability-level of the designated class?

(Circle One)

High ability 3 *37/*
Average ability 2
Low ability 1

11. How many students are enrolled in the designated class?

STUDENTS ☐☐ *38-39/*

12. Are all students in the designated class compiling mathematics portfolios?

(Circle One)

Yes 1 *40/*
No 0

IF NOT, which types of students are excused from preparing portfolios and why?

_____

_____

_____

13. If the designated class contains 8th grade students, which best describes the curriculum for the class?

(Circle One)

General mathematics 4 *41/*
Pre-algebra 3
Algebra I 2
Other mathematics 1
Describe _____

Designated class does not contain 8th grade students 0

CARD 01

Time Spent on Math Portfolios and Best Pieces

14. During a **typical month** how much **teacher time** (both in and out of class) is spent on the following aspects of the mathematics portfolios?

	Hours Per Month	
Finding appropriate tasks and/or materials	☐☐☐	42-44/
Preparing portfolio lessons	☐☐☐	45-47/
Conducting portfolio lessons.....	☐☐☐	48-50/
Helping students organize/manage their portfolios	☐☐☐	51-53/
Scoring/evaluating the contents of the math portfolios.....	☐☐☐	54-56/

15. During a **typical month** how much **student class time** is spent on the following aspects of the mathematics portfolios?

	Hours Per Month	
Doing portfolio tasks for the first time	☐☐☐	57-59/
Revising or rewriting portfolio tasks	☐☐☐	60-62/
Organizing/managing portfolios	☐☐☐	63-65/

16. Who selects the "best pieces" in your students' mathematics portfolios?

(Circle One)

Students select their own..... 3 66/

Students and teacher select in consultation..... 2

Teacher selects..... 1

17. What bases are used to select best pieces?

_____

_____

_____

Training and Support

18. Have you received training or assistance with the mathematics portfolios from any of the following sources?

(Circle One Number on Each Line)

	Yes	No	
Summer math institute	1	0	67/
Fall math institute	1	0	68/
Preparation-for-scoring workshop	1	0	69/
Network training activities at your school	1	0	70/
Informal training from:			
Other local teachers	1	0	71/
Your school principal	1	0	72/

19. Considered by itself, how well did each type of training or assistance prepare you to work with the mathematics portfolios?

(Circle One Number on Each Line)

	Very Well	Well	Adequately	Poorly	Very Poorly	Did Not Receive	
Summer math institute	5	4	3	2	1	0	73/
Fall math institute	5	4	3	2	1	0	74/
Preparation-for-scoring workshop	5	4	3	2	1	0	75/
Network training	5	4	3	2	1	0	76/
Informal training from:							
Other local teachers.....	5	4	3	2	1	0	77/
Your school principal	5	4	3	2	1	0	78/

20. Overall, how well did the available training and support prepare you to work with the mathematics portfolios?

(Circle One)

Very well	5	4/
Well	4	
Adequately	3	
Poorly	2	
Very poorly	1	

Changes in Mathematics Curriculum and Instruction

21. ***Prior to the portfolio assessment***, did you ever teach math to students at the same grade level as the designated class?

(Circle One)

Yes	1	→	*(Continue with Q22)* 5/
No	0	→	*(Skip to Q27)*

22. ***Comparing this school year to the last school year before you did portfolios***, how much change have you made in the amount of class time devoted to the following mathematical topics?

	CHANGE IN TIME					
	Much More	Somewhat More	About The Same	Somewhat Less	Much Less	
Computation and algorithms	5	4	3	2	1	6/
Estimation	5	4	3	2	1	7/
Patterns/Relationships	5	4	3	2	1	8/
Measurement/Geometry	5	4	3	2	1	9/
Problem Solving Strategies	5	4	3	2	1	10/

23. If you are spending less time this year on any other aspects of mathematics, please list them.

I am spending less time on:

_____

_____

_____

_____

24. ***Comparing this school year to the last school year before you did portfolios,*** how much has your emphasis on each of these aspects of **mathematical communication** changed?

	CHANGE IN EMPHASIS					
	Much More	Somewhat More	About The Same	Somewhat Less	Much Less	
Describing personal experiences ..	5	4	3	2	1	*11/*
Oral discussions of mathematics...	5	4	3	2	1	*12/*
Making charts, graphs, diagrams, etc.	5	4	3	2	1	*13/*
Written reports about mathematics	5	4	3	2	1	*14/*

25. ***Comparing this school year to the last school year before you did portfolios,*** how much time do students spend on the following **problem solving activities**?

	CHANGE IN EMPHASIS					
	Much More	Somewhat More	About The Same	Somewhat Less	Much Less	
Exploring mathematical patterns .	5	4	3	2	1	*15/*
Single-step word problems	5	4	3	2	1	*16/*
Multiple-step word problems	5	4	3	2	1	*17/*
Logic or reasoning problems	5	4	3	2	1	*18/*
Applying math knowledge to new situations	5	4	3	2	1	*19/*
Collecting, analysing, reporting data.....	5	4	3	2	1	*20/*

26. ***Comparing this school year to the last school year before you did portfolios,*** have you changed the **pattern of student grouping** you use for mathematics instruction?

	CHANGE IN TIME					
	Much More	Somewhat More	About The Same	Somewhat Less	Much Less	
Whole class instruction	5	4	3	2	1	*21/*
Small groups working together	5	4	3	2	1	*22/*
Pairs of students working together	5	4	3	2	1	*23/*
Individual work	5	4	3	2	1	*24/*

27. How often do you assign the following types of problems in the designated class?

	APPROXIMATELY						
	Every Lesson	2-3 Per Week	Once Per Week	1-3 Per Month	1-3 Per Term	Never	
Puzzles (non-routine problems dealing with logic and reasoning) ..	5	4	3	2	1	0	25/
Applications (situational applications of knowledge students already possess).....	5	4	3	2	1	0	26/
Investigations (exploration, data collection and analysis)	5	4	3	2	1	0	27/

Portfolio Preparation

28. How much emphasis do you place on making sure the individual pieces in students' math portfolios have each of the following characteristics?

	AMOUNT OF EMPHASIS				
	Heavy	Moderate	Minor	None	
Mathematically correct and complete.....	3	2	1	0	28/
Neat and polished in appearance.....	3	2	1	0	29/
Interesting or important to students	3	2	1	0	30/
Similar to examples in the Resource Guide	3	2	1	0	31/

29. On average, how many times do students in the designated class rewrite or revise the work they place in their math portfolios?

TIMES ☐☐ 32-33/

30. What kinds of limits do you place on the amount of help students can obtain from parents or other adults when completing portfolio projects?

(Circle One)

No limits 4 34/

Limits on the amount of time adults can spend helping students..... 3

Limits on the type of assistance adults can give to students..... 2

Limits on both the amount of time and the type of assistance..... 1

31. Are your students required to acknowledge assistance from parents or other adults on their portfolio pieces?

(Circle One)

No acknowledgement is required..... 3 *35/*

Students must indicate that they had assistance, but need not describe it..... 2

Students must describe the kind or amount of assistance they received..... 1

32. How often do you use the following sources to find portfolio problems or tasks for the designated class?

(Circle One Number on Each Line)

	Very Frequently	Regularly	Occasionally	Seldom	Never	
Vermont mathematics portfolio Resource Book	4	3	2	1	0	*36/*
Handouts from summer or fall institutes	4	3	2	1	0	*37/*
Materials shared by Network leader or teachers	4	3	2	1	0	*38/*
Materials or ideas from teachers at your school	4	3	2	1	0	*39/*
Tasks you made up yourself	4	3	2	1	0	*40/*
Tasks from textbooks or other published books	4	3	2	1	0	*41/*
Other sources	4	3	2	1	0	*42/*

Student Performance

33. Have you changed your opinion of students' mathematics abilities based on their portfolio work?

(Circle One)

A great deal..... 4 *43/*

A moderate amount..... 3

A little bit..... 2

Not at all..... 1

CARD 02

34. Do students perform differently on portfolio tasks than on other mathematics work?

(Circle One)

Generally better on portfolio tasks	4	44/
About the same on both kinds of assignments.....	3	
Generally worse on portfolio tasks	2	
Performance varies too much across students and/or tasks to make a general statement.....	1	

Evaluation of Portfolio Assessment

35. What positive effects do the mathematics portfolios have on you and your students? Please indicate how often the following statements are true for you and for the designated class this year.

(Circle One Number on Each Line)

	Always	Often	Occasionally	Rarely or Never	
I am more enthusiastic about teaching math.....	4	3	2	1	45/
Goals of math instruction are improved	4	3	2	1	46/
Math is more closely linked to other subjects	4	3	2	1	47/
Students' attitudes toward math improve.....	4	3	2	1	48/
Students are learning more mathematics	4	3	2	1	49/
Low-ability students are more successful	4	3	2	1	50/

CARD 02

36. What problems do the mathematics portfolios create for you and your students? Please indicate how often the following statements about the mathematics portfolios are true for you and for the designated class this year.

(Circle One Number on Each Line)

	Always	Often	Occasionally	Rarely or Never	
I don't understanding <u>what</u> I am expected to do	4	3	2	1	*51/*
I don't have enough training in <u>how</u> to do it.....	4	3	2	1	*52/*
I have difficulty finding appropriate tasks	4	3	2	1	*53/*
I lack time to prepare portfolio lessons	4	3	2	1	*54/*
There is not enough time to cover the full math curriculum	4	3	2	1	*55/*
Low ability students have difficulty with tasks ..	4	3	2	1	*56/*
Students don't understand <u>what</u> to do with tasks	4	3	2	1	*57/*
Students don't know <u>how</u> to solve problems.....	4	3	2	1	*58/*
Students are not interested in portfolio tasks....	4	3	2	1	*59/*

Thank you very much for your cooperation.

CARD 02